COMPETING RISKS MODELS OF ECONOMIC BEHAVIOR:

THEORY AND APPLICATIONS TO RETIREMENT AND UNEMPLOYMENT

by

GLENN TETSUMI SUEYOSHI

B.A., Pitzer College (1983)

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1987

Signature of Author ______________________________
Department of Economics

Certified by ______________________________
Jerry A. Hausman, Thesis Supervisor

Certified by ______________________________
James M. Poterba, Thesis Supervisor

Accepted by ______________________________
Peter Temin
Chairperson, Departmental Graduate Committee

Room 14-0551
77 Massachusetts Avenue
Cambridge, MA 02139
Ph: 617.253.2800
Email: docs@mit.edu
http://libraries.mit.edu/docs

Competing Risks Models of Economic Behavior: Theory and Applications to Retirement and Unemployment

by

Glenn Tetsumi Sueyoshi

Submitted to the Department of Economics
on September 1, 1987 in partial fulfillment
of the requirements for the Degree of
Doctor of Philosophy

Abstract

Economists increasingly have become involved in the development and use of statistical models for the analysis of duration data. This thesis consists of three chapters which relate to the specification and application of these models to the study of economic phenomena.

In chapter 1, I analyze the determinants of retirement in a competing risk framework, and find that distinguishing between full and partial retirement is important if one wishes to understand the factors influencing the decision to retire. In particular, the Social Security system affects full and partial retirement probabilitites differentially. Chapter 2 consists of an analysis of the duration of unemployment spells. Using a newly developed data source I find that there is mixed support for a simple search interpretation of spell duration. The evidence suggests that a more general framework is required. In chapter 3, I extend the regression form of the proportional hazards model to the case where covariates are allowed to vary over time. I demonstrate identification and asymptotic normality of the general estimator for both single and competing risks models.

The results in this thesis address specific questions regarding Social Security and unemployment insurance as well as issues relating to the general use of duration models in explaining economic phenomena. The framework of analysis used in this thesis should profitably extend to a number of areas of future research.

Thesis Supervisors: Jerry A. Hausman, Professor of Economics
James M. Poterba, Associate Professor of Economics

Table of Contents

Acknowledgements 4

Introduction. 6

Chapter 1: Social Security and the Determinants of Full and Partial Retirement: A Competing Risks Analysis

Introduction 10

1. Background 13
   1.1 Social Security. 13
   1.2 Partial Retirement 20
      1.2.1 The Definition of Retirement. 20
      1.2.2 The Extent of Partial Retirement. 23

2. The Existing Literature. 31

3. Econometric Specification. 36
   3.1 Single Risk Estimation 37
   3.2 Competing Risks Estimation 43

4. Data 48

5. Results. 53
   5.1 Single Form of Retirement. 56
      5.1.1. Parametric Baseline Hazard (Weibull) . . . 56
      5.1.2. Semi-Parametric Baseline Hazard. 61
   5.2 Full and Partial Retirement. 75
      5.2.1. Independent Risks. 75
      5.2.2. Correlated Risks 84

Conclusion 93

Chapter 2: Unemployment Insurance Benefits and the Duration of Unemployment Spells: Evidence from the Survey of Income and Program Participation

Introduction 95

1. Background 98

2. Data Issues. 102

3. Results. 109
3.1 Descriptive Results. 109
3.2 Single Risk Models 128
3.3 Competing Risks Models 139

4. Conclusion 152

Chapter 3: Semi-parametric Proportional Hazards Estimation of Single and Competing Risks Models with Time-varying Covariates

Introduction 154
1. Single Risk Duration Models. 157
1.1 Specification. 157
1.2 Identification, Consistency and Asymptotic Normality. .163

2. Competing Risks Duration Models. 176
2.1 Specification. 176
2.2 Identification, Consistency and Asymptotic Normality. .180

3. Conclusion 193

Appendix 1. 195

Appendix 2. 198

Appendix 3. 201

References. 203

Acknowledgements

I would, as is customary, like to acknowledge the contributions of those without whom this thesis could not possibly have been completed.

First and foremost, I wish to thank my thesis advisors, Jerry Hausman and James Poterba, for wading through my excessively lengthy prose and providing valuable direction and comments. Perhaps more importantly, they tolerated my peculiar last-minute working style and I am particularly indebted to them for the rapid turnaround time on my manuscripts. My thanks also go to Aaron Han and Hidehiko Ichimura for help in clarifying a variety of econometric issues associated with duration models and to Steve Venti for timely provision of the SIPP data.

One of the externalities associated with graduate study at M.I.T. is the value of the contributions that one's colleagues make to the final version of a thesis. Rather than provide a lengthly list of individuals, let me instead gratefully acknowledge the immense collective contribution of the M.I.T. economics community to this thesis. I am also grateful for the financial support provided by the National Science Foundation and M.I.T. economics department.

Having good colleagues who are also good friends is more than one should expect and in that regard, I consider myself quite fortunate. At the risk of offending those excluded, I would particularly like to thank James Dana, Janice Eberly, William

English, Robert Gertner, Hidehiko Ichimura, Gregory Leonard, Pamela Loprest, Raymond McFadden, Anne Maasland, Bradley Reiff, Danny Quah, Christopher Vellturo, and especially Leslie Papke, for making graduate student life bearable.

Above all, my family deserves a lot of credit for putting up with my particular idiosyncracies for the past 25 years. They were never quite certain what I was doing with my life, but trusted me enough to find my own direction. And lastly, I cannot overstate the importance of Julie Boyer who has, in inumerable ways and in her inimitable style, been an integral part of the past seven years of my life. Boyer has always been there when needed, and her importance is immeasurable. To my family and Julie, this thesis is respectfully dedicated.

Introduction

Economists increasingly have become involved in the development and use of techniques for the analysis of duration data. Duration, or hazard, models provide a convenient framework for examining economic phenomena which relate to the length of time before an event occurs. This thesis consists of three chapters which involve the use of hazard models in economics.

The first two chapters are comprised of empirical analyses using semi-parametric techniques for the estimation of competing risks models. In the first chapter, I analyze the factors influencing the number of years to retirement and the form, full or partial, that the intitial retirement takes. The second chapter considers the factors that affect unemployment spell durations and whether exit is via recall or a new job finding. In the third chapter, I demonstrate identification and asymptotic normality for an estimator which generalizes previous semi-parametric estimators to allow for explanatory variables which change over time. The development of this estimator continues the effort, initiated by others, directed at finding a unified framework for the analysis of duration data.

In chapter 1 I examine the relationship between Social Security and retirement behavior. Empirical analyses of retirement typically assume a single form of retirement. There is evidence, however, that a substantial proportion of individuals exit from full time work via a partial reduction in work effort. To the extent that behavior differs across retirement type, then single form of retirement models are likely to convolute the influences of various factors upon the two forms of retirement. If, for example, the receipt of Social Security benefits has differential effects by type of retirement, then the single form model will not accurately capture the impact of benefits upon behavior.

To account for the existence of partial retirement, I use a competing risks model of full and partial retirement to analyze retirement decisions. The individual level data are drawn from a sample of Longitudinal Retirement History Survey individuals. Based upon parameter estimates from this specification, I find evidence that single form of retirement models convolute the influence of variables. Additional Social Security benefits are found to increase the probabilities of retirement differentially across retirement type, encouraging full retirement more than partial retirement. A similar result is observed for increases in benefit levels resulting from additional work which lower the relative probability of partial retirement.

The second chapter addresses the impact of unemployment insurance upon the duration of spells of unemployment. Despite considerable advance in both economic theory and econometric technique, much of the existing empirical work on the subject has been handicapped by data limitations which hamper the study of a variety of issues. In this chapter, I consider jointly the issue of the impact of UI benefits upon exit from spells of unemployment and the form in which the exit occurs. The primary contribution of the chapter is the development of a new sample of unemployment spells from the Survey of Income and Program Participation which allows for the estimation of a competing risk model of recall or new job finding and which addresses a number of data problems.

I find that unemployment benefits and the exhaustion of benefits are important factors in influencing hazard rates. There is mixed support in the data for a search model of new job finding. While the overall hazard is estimated to be a decreasing function of duration, the new job hazard for exit from unemployment appears to be increasing over time, but not monotonically. The relative importance of unemployment insurance in affecting the hazards for recall is a result that is beyond the scope of search models to explain and suggests the need for more general models of unemployment durations. These results, along with the finding of a large number of multiple spell individuals, indicate that the recent concern with the duration of single spell unemployment is somewhat missplaced.

The final chapter builds upon the previous literature on semi-parametric hazard models by extending the regression form of the proportional hazards model to time-varying covariates. Recent econometric research on proportional hazards models has focused on estimation methods which do not require functional form restrictions on the form of the baseline hazard. Unfortunately, those techniques based upon the regression form of the likelihood have, to date, been restricted to the special case of covariates which do not vary over time.

I demonstrate identification and asymptotic normality of the general estimator for single and competing risks models with time-varying covariates. The proofs borrow heavily from the existing literature on hazard models and discrete choice estimators. In the course of demonstrating the properties of the estimator, I take advantage of and point out the obvious correspondence between the semi-parametric estimation techniques and existing discrete choice models. This correspondence should provide the basis for future research on the estimation of hazard models.

Taken together, the three chapters of this thesis address a broad range of issues relating to the specific application of duration models to the study of retirement and unemployment, and the general use of these models in explaining a variety of economic issues. The framework of analysis used in this thesis should profitably extend to a number of important areas of current concern.

Chapter 1

Social Security and the Determinants of Full and Partial Retirement: A Competing Risks Analysis

Introduction

The Social Security system affects individual intertemporal budget constraints through a complicated combination of taxes and transfers. Despite the complexity of the task, several researchers have been able to document the most important of these effects, tracing the impact of a variety of Social Security rules upon the budget constraint of a given individual.[1] However, the labor supply effects of changes in these constraints are less well understood. Despite a number of attempts at quantification, important unanswered questions remain regarding the relationship between the Social Security system and retirement behavior. In this chapter I consider a particular variant of this theme--the interrelated questions of what factors influence the decision to retire and the form that the retirement will take.

This chapter extends the previous analysis on the retirement question in several ways. First, I estimate a model of retirement in

---

[1]See Blinder, Gordon and Wise [1980] and Aaron [1984] for relatively comprehensive discussions of the issues involved.

which partial retirement is treated distinctly from full retirement. To date, most studies that have considered the retirement question have relied on a single measure of retirement.[2] The use of a single measure has the effect of confounding the different motivations lying behind the two decisions and leads to misleading inferences about the effects of variables upon retirement decisions. The estimation of a model which differentiates between full and partial retirement also allows one readily to consider the effects of policy changes upon the relative frequencies of the two forms of retirement.

This is also, to the best of my knowledge, the first analysis of retirement behavior that uses a competing risks framework to analyze full and partial retirement. In this chapter I use new econometric techniques developed by Han and Hausman [1986] for the study of duration data which allow for relatively few functional form restrictions and the potential for correlation between retirement risks.

The use of competing risks has several advantages for the study of retirement behavior. For one, the use of duration models is natural given the dynamic nature of the retirement decision. These models allow results to be expressed in terms of changes over time in the probabilities of retirement. Second, duration models allow the

---

[2]Notable exceptions include Boskin [1977], Gustman and Steinmeier [1984 and 1986] and Zabalza *et.al.* [1980]. In addition, Burtless and Moffitt [1984] consider post-retirement work behavior.

treatment of changing predetermined variables which most static models must assume away.[3] Finally, the notion of a hazard rate, or conditional probability of retirement, accords with the way in which most individuals think about retirement in that they condition their retirement decisions anew each period given the information they have at that time.

Using these econometric techniques, I find that partial retirement behavior differs substantively from full retirement behavior. The dissimilarity of the two forms of retirement is reflected in the differing effects of Social Security upon the probabilities of retirement. In particular, I find that partial retirement is more strongly influenced by economic variables than is the corresponding full retirement decision. Furthermore, it appears that full retirement may be strongly motivated by factors such as health and occupation that are outside the realm of traditional economic incentives.

In these results, Social Security has significant effects upon retirement through both benefit levels and potential increases in benefits. The finding that changes in the Social Security system affect the two forms of retirement in different ways should be of

---

[3] In this chapter I do not consider the effects of time-varying covariates upon the retirement process. The theoretical results of chapter 3 of this thesis extend the semi-parametric techniques of Han and Hausman to allow for changes over time in the predetermined variables. I plan to implement this estimator in subsequent work.

interest to policymakers. Simulation results indicate that the early 1970s growth in the Social Security system has increased the probability of full retirement and reduced the probability of partial retirement. The changes in probabilities account for a portion of the observed decline in labor force participation, but probably cannot be viewed as the primary factor.

The remainder of this chapter is organized as follows. Section 1 provides an overview of the basic questions regarding the nature and causes of retirement. I also discuss the institutional features of the Social Security system that play a role in the retirement decision and highlight the importance of partial retirement as a means of exit from full-time employment. In section 2, I review briefly the existing empirical literature. The econometric specification is outlined in section 3, and the data are described in section 4. The empirical results from the estimation of duration models of retirement are presented in section 5. In addition to parameter estimates, I discuss the results of simulations designed to analyze the effects of recent changes in Social Security law upon retirement behavior. There is a concluding section.

1. Background

1.1 Social Security

Many individuals have expressed a general concern with the possible work-disincentive effects of Social Security, citing a correlation between trends in elderly labor force participation rates and Social Security benefit and wealth levels. In table 1, legislated benefit increases are traced over the two decades of rapid growth from 1960-1980. The growth in benefits was particularly rapid in the early 1970s, with real primary insurance benefits increasing by a minimum of 19.4 percent. Table 2 presents corresponding statistics showing the decline in labor force participation rates for both males and females over roughly the same period. Again, there are relatively large changes in the early 1970s. Between 1970 and 1975, the participation rate for males aged 55-64 fell from 83.0 to 75.8, a decrease almost 5 times larger than that experienced over the preceeding 5 years.[4] These figures on benefit levels and labor force participation have generated considerable empirical interest in the question of how much of the decline in participation rates can be attributed to the growth of Social Security.

There is some theoretical justification for concern over the relationship between Social Security and early retirement. For individuals between the ages of 62 and 72 considering whether to continue working or to retire, the Social Security system creates significant non-linearities in the return to work. In essence,

---

[4]The change in the rates for females between 1970 and 1975 shows a similar increase relative to the preceeding 5-year period, though the magnitudes and directions of the changes over the 20-year period are clearly influenced by a variety of other social factors.

Table 1 - Cumulative effect of statutory and automatic increases in real primary insurance benefits: minimum percentages, 1959-80 -- select dates.

	Date of Comparison			
Base Date	Jan 1965	Jan 1970	June 1975	June 1980
Jan 1959	-1.2	9.2	35.3	48.3
Jan 1965		6.9	11.4	48.8
Jan 1970			19.4	26.8
June 1975				-2.1

Source: Social Security Administration, *Annual Statistical Supplement* (1981), p. 27.

Table 2 - Labor force participation rates for the elderly, by age and sex: 1960-1980

	1960	1965	1970	1975	1980
Male					
55-64	86.8	84.6	83.0	75.8	72.3
65 +	33.1	27.9	26.8	21.7	19.1
Female					
55-64	37.2	41.1	43.0	41.0	41.5
65 +	10.8	10.0	9.7	8.3	8.1

Source: Bureau of the Census. *Statistical Abstract of the United States 1981*, p. 381.

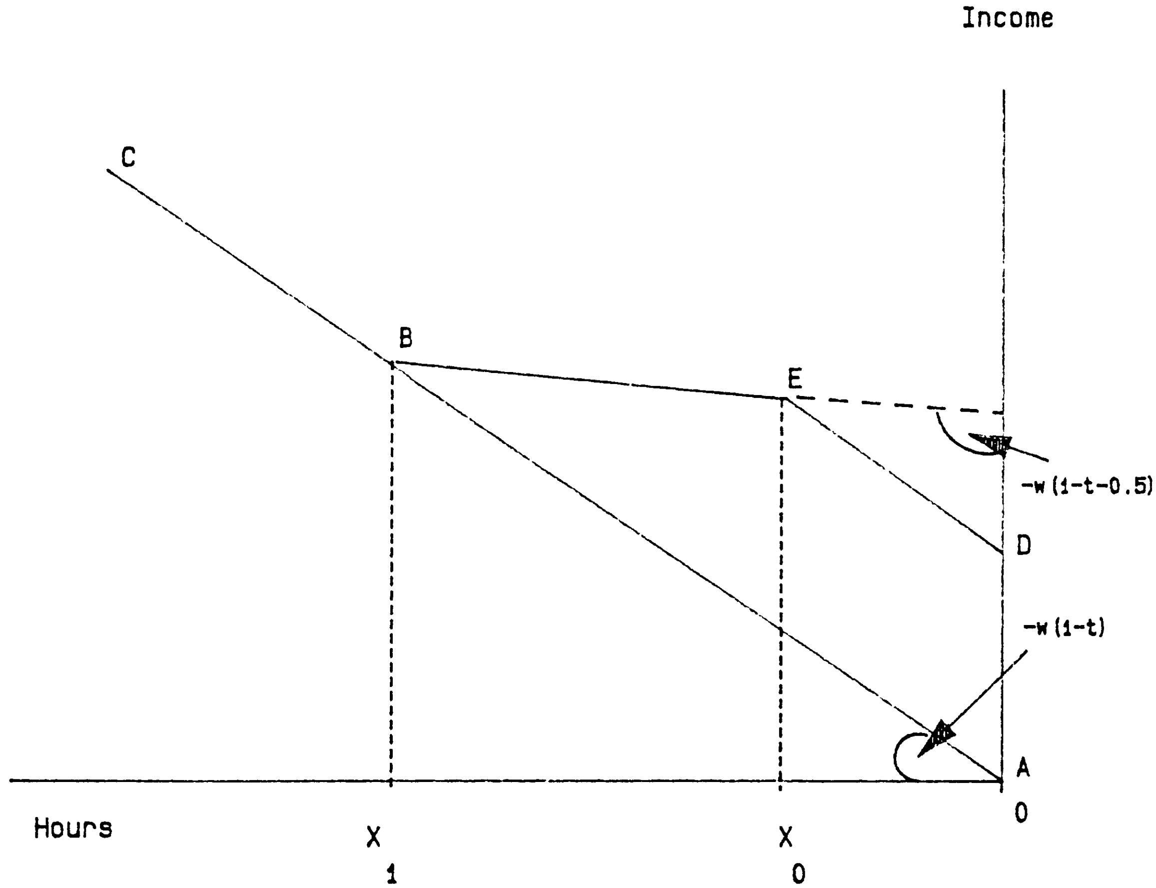
beginning at age 62, the basic benefit level is taxed at a 50% rate for earnings over a specified level until the benefits are exhausted.[5] The tax schedule yields the familiar non-linear budget constraint depicted in figure 1. Prior to age 62, an individual faces the budget set *ABC* which is governed by the after tax wage $w(1-t)$. At age 62, the individual is eligible to receive a basic benefit level of B if he retires. Up to an earnings level of X_o, the marginal incentives are unchanged so that the slope of the budget line is given by the negative of the earlier after-tax wage. At that point, the earnings test taxes the basic benefit at a 1:2 rate so that the slope of the budget line is now $w(1-t-0.5)$. At X_1, the basic benefits are exhausted and the individual's budget set lies along the *BC* portion of the original segment. The total budget set under the Social Security system is thus *DEBC*.

Several analysts have seized upon the implications of figure 1 as a reason for concern with the work-incentive effects of the Social Security system. The unambiguous conclusion that Social Security reduces labor force participation does not, however, follow immediately. Drawing inferences based upon figure 1 alone can be misleading since the diagram does not fully capture the relevant incentives. In particular it ignores the life-cycle nature of the choices involved as well as salient features of the Social Security

[5]In 1981, the earnings test allowed a maximum of $6,600 in earnings before benefits were reduced.

Figure 1 - Non-Linear Budget Constraint For Individual

Eligible for Social Security Benefits



system.

First, the existence of automatic benefit recomputation (ABR) implies that the return to work consists of increments to Social Security wealth in addition to earnings. The effects of ABR have been outlined in some depth by Blinder, Gordon and Wise (BGW) [1980]. In brief, the Social Security benefits that are received by an individual upon retirement are calculated on the basis of the lifetime work history. The first step involves calculating a measure of the individual's average monthly wages (AMW) which, to simplify somewhat, are computed by averaging over the T highest years of earnings in covered employment.[6] The primary insurance amount (PIA) is obtained by applying the AMW to a progressive benefit formula. An unmarried individual retiring at the base age of 65 receives the PIA; a married couple receives either the sum of the individual entitlements, or 1.5 times the single benefit level, whichever is greater.[7] An individual who works for an additional year is able to

---

[6]The value of T depends upon the age of the individual and the year in which the computation takes place. For the individuals in the RHS, the relevant formula is

$$T = \min(\text{year},\text{yr65}) - 1956$$

where year = the year of the computation and yr65 = the year the individual turns 65. For currently retiring individuals, the number of years of computation is the minimum of the current age less 26, and 36.

[7]Beginning in 1974, benefits were indexed for inflation. The PIA is multiplied by the ratio of the growth in the CPI over the interval from 1974 to the year benefits are received. As Diamond [1977] notes, this creates significant overindexation for inflation since

substitute current earnings for the minimum of the earnings currently used in the computation of the AMW. Individuals with upward sloping age-earnings profiles are likely to have much higher current earnings than those received in the lowest earnings period, especially in an inflationary environment, so that the increases in the AMW and hence in future Social Security benefits can be significant. BGW perform calculations which suggest that under pre-1977 law, the wealth effect is on the order of a 50% wage subsidy for a representative individual.

Second, the effects of the imperfect actuarial adjustment mean that the relevant lifetime budget constraint can be improved by additional work up through age 65. Actuarial adjustments are made to the basic benefit with the stated purpose of providing statistically fair increases and decreases in benefits for those accelerating or postponing retirement. Individuals who elect to retire prior to age 65 have their monthly benefits decreased to account for the longer expected period of time over which they will be receiving those benefits. The actuarial reduction is 5/9% for each month between ages 62 and 65 in which the individual does not draw benefits. No benefits are paid prior to age 62. Similarly, benefits are increased

---

the indexation is from a fixed date rather than from the date the individual turns 62, and since, prior to 1977, the AMW computations were based upon nominal income. From 1977 onward, the AMW is computed using earnings indexed by the growth in overall earnings levels.

by 1/12% for each month without payment of benefits between ages 65 and 72.[8] Calculations made by BGW indicate that these actuarial adjustments are fair or more than fair for the 62-65 period and less than fair for the post-65 period. Thus, there is an actuarial bonus to delaying retirement to age 65, but a penalty for delaying past 65.

The effects of ABR and actuarial adjustments are difficult to capture with a single-period model of this type, but if thought of as wage subsidies, can be represented by increases in the slope of the budget line under Social Security. For large enough wage subsidies, the income and substitution effects operate in opposite directions, with the substitution effect encouraging work. The result of the various income and substitution effects are such that the exact relationship between Social Security and retirement is theoretically ambiguous. Furthermore, assessing the magnitude of the effects of Social Security upon retirement requires empirical analysis.

1.2 Partial Retirement

1.2.1 The Definition of Retirement

There is no natural way in which to define retirement, consequently a variety of definitions for retirement have been used in the empirical literature. Among them are complete withdrawl from the labor force (Gordon and Blinder [1980]), transition from job held

---

[8] Recent legislation has allowed the full payment of benefits once an individual reaches age 70, regardless of employment status.

at the start of the sample (Fields and Mitchell [1984]), self-reported status (Hausman and Wise [1985]), the receipt of pension income (Burkhauser [1979]), and work less than a specified number of hours (Boskin [1977]).

Difficulties arise with the use of each of these definitions. Under the definition based upon labor force participation, individuals working a few hours a week but who are substantively retired and perhaps drawing pension benefits will be classified as working. Under a transition definition, individuals who make job changes totally unrelated to retirement will be classified as retired. Self-reported status suffers from being unrelated to observable economic activity. Available evidence suggests that using pension receipt as an indicator of retirement may be misleading because many individuals who are not working at all receive no pension income (Diamond and Hausman [1984]). Finally, there is a certain arbitrariness to definitions based upon reduced work effort. What is most disturbing is the result that these definitions often yield conflicting classifications.[9]

Underlying these definitional difficulties is the fact that a single form of retirement is unable to capture the behavior of individuals who are not employed in a "standard" full-time job, but

---

[9]To see the problems involved, consider the case of an individual who reduces the number of hours worked per week in the same job by a third in order to collect a newly vested partial pension.

are still in the labor force. It is this fundamental problem that motivates the consideration of partial retirement in this paper.

Like full retirement, partial retirement can be defined in a number of ways, but central to the concept is a discontinuous, though not complete, reduction in work effort. An individual who wishes to reduce his work effort can do so by reducing the number of hours worked per week, by reducing the number of weeks worked per year, by working less diligently during the time spent at work, or through some combination of the above. For the purposes of this study, partial retirement is defined as employment at a job in which the individual reports less than 35 hours of work per week or employment for less than 46 weeks (the weeks-equivalent of a 35 hour work week).[10] The definition corresponds to considering reductions in the hours and weeks worked, but ignores the unobservable measure of work effort.[11]

In the subsequent discussion, retirement of either form will be

---

[10]Weeks of unemployment are considered as weeks of employment when evaluating the latter measure. It is possible that individuals will report vacation time as weeks of no work so that the weeks definition will overstate the extent of partial retirement. This does not appear to be a problem since almost all individuals who are classified as partially retired under the weeks definition are also partially retired by an hours criterion.

[11]This definition suffers from the arbitrariness cited above. Among the studies that have considered partial retirement, Boskin [1977] and Zabalza *et. al.* [1980] use reduced hours of work; Gustman and Steinmeier [1986] use self-reported status. The particular choice of definition used in this study is motivated in part by the questioning patterns of the Longitudinal Retirement History Survey, and in part by a desire to base the definition on observed behavior.

referred to simply as retirement; where distinguishing between full and partial retirement is necessary, the appropriate qualifier will be used.

1.2.2 The Extent of Partial Retirement

Typically, analyses of retirement behavior have not considered the possibility of partial retirement and have defined a single form of retirement in one of the ways listed above. The notion that partial retirement is distinct from full retirement and should be treated separately in empirical work was first adopted by Boskin [1977] and has been used by, among others, Boskin and Hurd [1978] and Zabalza *et. al*. [1980]. For the most part, however, researchers have utilized the concept of a single form of retirement.[12]

Recent work by Gustman and Steinmeier [1984] has shed some light on the extent of partial retirement. Using a sample drawn from the Longitudinal Retirement History Survey (RHS), they find that of those individuals who are observed to transit from full-time employment, approximately 28.2% partially retire. Over one-third of the individuals studied report partial retirement in at least one of the four sample periods. These results are corroborated by Zabalza *et. al.* who report that their sample of elderly in Great Britain exhibits a well-defined bimodal distribution of hours worked, and by Parnes

[12] Relatively comprehensive surveys of the literature are provided by Fields and Mitchell [1984] and Aaron [1984].

and Nestel [1981] who find that about 20% of their 1966-1975 National Longitudinal Survey sample reports post-retirement labor market activity.[13] Moreover, given current and anticipated future reductions in mortality rates and the resulting increases in longevity, the importance of partial retirees in the economy is likely to grow over time.

Drawing upon data derived from the RHS, I find results that are similar to those of Gustman and Steinmeier, despite the use of an hours/weeks worked definition rather than self-reported status. Table 3 shows the number of individuals retiring at various ages, aggregated across cohorts, broken down by type of retirement. The censoring category refers to those individuals who do not retire over the sample interval. The age at censoring is therefore the last age at which they were observed. The sample includes only those individuals who retire after age 57. Since less than 1/2 of one percent of the individuals report retirement prior to age 58, and of those individuals, one-quarter appear to be subject to coding error, or are highly unrepresentative of the population as a whole, this

---

[13] In the latter study, retirement is defined by self-reported status. Among those classified as retired who reported having worked in the 12 month period prior to the 1976 survey, approximately 17 percent reported working more than 2,000 hours. Assuming work over the entire year, this translates into about 38.5 hours per week. 23 percent of the sample reported working more than 29 hours per week, 42 percent more than 19 hours per weekk,and 71 more than 10 hours per week.

Table 3 - Age distribution of retirement or censoring by type of first retirement, 1633 person RHS subsample.

Age at Retirement	Type of Initial Retirement (Number of Persons)			
	Full	Partial	Censored	Total
58	4	1	0	5
59	6	0	0	6
60	13	11	7	24
61	27	14	6	41
62	86	61	7	147
63	124	56	7	180
64	117	130	4	247
65	275	75	8	350
66	173	65	1	238
67	67	52	1	119
68	36	20	29	56
69	19	16	23	35
70	14	5	16	19
71	5	6	14	11
72	7	1	12	8
73	2	2	8	4
Totals	975	515	143	1633

Note: Based upon author's calculations. The various forms of retirement are defined in the text.

procedure should not create undue biases.[14] Focusing first on the number of individuals in each retirement class, table 3 shows that out of a sample of 1633 individuals, over 500 partially retire. Since 975 individuals move directly to full retirement, a little under a third of those who actually retire over the sample period partially retire. While a bit higher than the figure found by Gustman and Steinmeier, the observed fraction is certainly comparable in magnitude.

In table 4 I present data on mean age of retirement for various types of retirement. To see the impact of differing definitions, note that if one defines retirement to be the first instance of full or partial retirement, the mean retirement age is 64.75; this average age is in contrast to a mean retirement age of 66.19 resulting from a definition based upon complete withdrawl from the labor force. Breaking down the mean age of initial retirement, the mean age for partial retirees is 64.68, slightly less than the corresponding age of 64.79 for full retirees.

The closeness of these mean values hides a considerable amount of work activity. For those individuals who partially retire and are

---

[14]The ages of retirement for these individuals include 36, 46 and 48. The calculations of the percentage who are excluded on the basis of retirement age are based upon preliminary work utilizing a slightly different subsample than the one used in the estimation. While it can be argued that this exclusion creates a form of self-selection which biases the results, it seems more likely that these individuals differ from the rest of the population in the way that retirement decisions are made and should therefore be considered separately.

Table 4 - Mean age of retirement under different retirement paths, RHS subsample.

	Mean	*Std. Dev.*	*N*
Mean Age of Initial Retirement	64.752	2.201	1490
By Type of Retirement			
Full	64.791	2.169	975
Partial	64.678	2.260	515
Mean Age of Final Retirement	66.186	3.151	1347
By Type of Initial Retirement			
Full	64.791	2.169	975
Partial	69.841	2.269	372
Mean Duration of Partial Retirement	5.500	2.037	372

Note: Based upon author's calculations. All samples are limited to completed spells for the retirement in question. The partial retirement category for mean age of final retirement refers to those individuals observed to fully retire who first partially retired.

subsequently observed to fully retire, the mean duration of partial retirement is approximately 5.5 years. As one might expect, the duration is negatively correlated (-0.446) with the age of initial retirement so that individuals who retire earlier remain in partial retirement for a longer period of time. The mean duration is in excess of the durations studied by Gustman and Steinmeier who find that most of those who partially retire spend a relatively short period of time, one to two years, in that state. The discrepancy results, no doubt, not only from the differences in the definitions of partial retirement, but also from the differing definitions of full retirement since I use a definition based upon complete withdrawl from the labor force while Gustman and Steinmeier use self-reported full retirement. The latter definition will classify as fully retired individuals listed as partially retired by my definition.

In addition to differences between mean retirement ages, there are differences in the distribution of retirement ages. Table 3 shows that the number of individuals classified as retiring at a given age rises slowly up to age 62, then sharply through age 65, declining in subsequent years. The same is true when considering either full or partial retirees alone, though the number of partial retirees is generally smaller than the corresponding number of full retirees. The modal age for those who partially retire is age 64, earlier than for those who first fully retire, a result that is reflected in the slightly lower mean age of retirement. Somewhat surprisingly, the total number of partial retirees at age 64 is greater than the corresponding number of full retirees. The relationship is reversed at virtually every other age.

More informative, perhaps, are simple calculations of the hazard rates for both types of retirement. These represent the probability of retiring at a given age conditional on not having retired prior to that age. Thus, the full retirement hazard for a 63 year old individual is the probability that he fully retires at age 63, given that he has not retired previously. The hazards are calculated by dividing the number of retirees of a given type, (full or partial) at a given age, by the number of unretired individuals up to that age.[15]

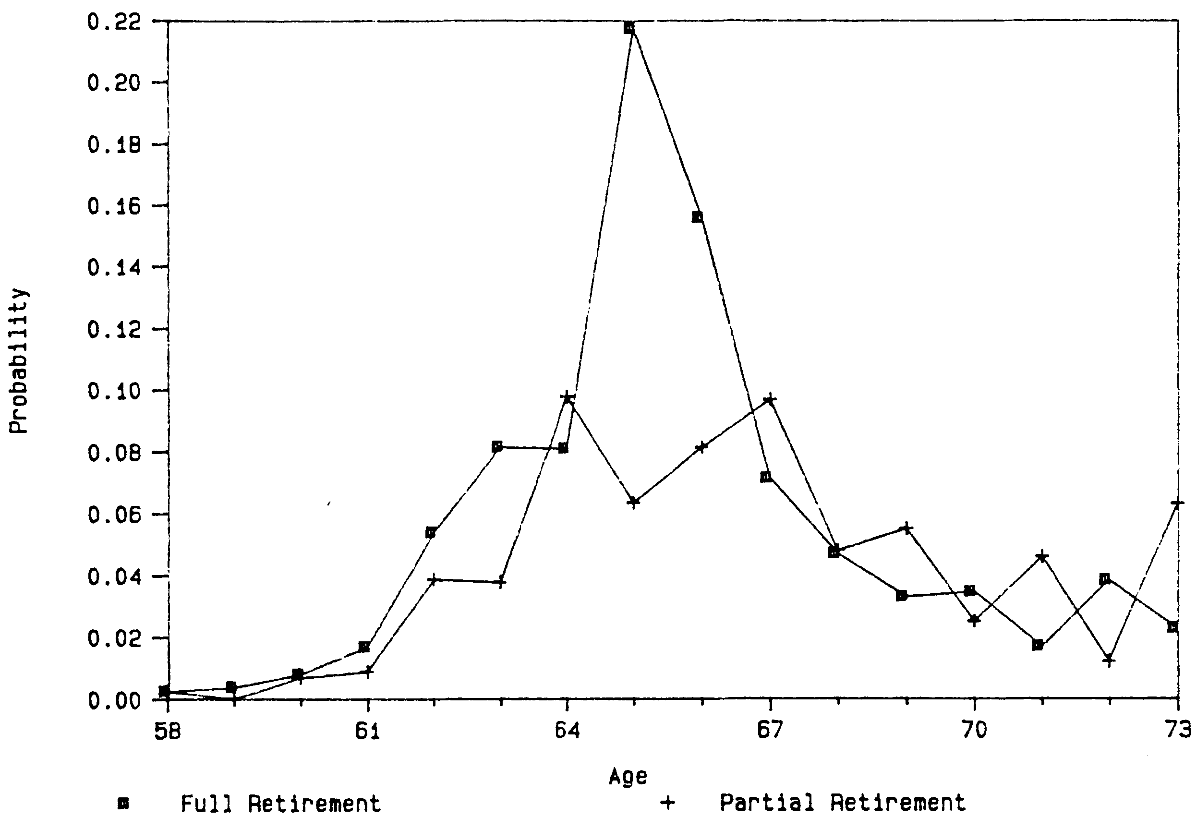
Hazard rates for both forms of retirement are depicted in figure 2. Up to age 64, the hazards are similar in shape, with the full retirement hazard rising somewhat faster than the partial retirement hazard. There is a pronounced spike in the full retirement hazard at age 65 that is not present for partial retirement. If anything, the partial retirement hazard is bimodal, with a pronounced drop in the hazard at age 65 and peaks at age 64 and 66. Moreover, with the exception of ages 65-66, the partial retirement hazard lies above the full retirement hazard for most ages greater than 64. The difference

---

[15] It is easy to show that this is the maximum likelihood estimator of the hazard rate for a homogeneous sample. This estimator is referred to in the statistics literature as the Kaplan-Meier estimator.

Figure 2 - Sample Hazard Rates for Retirement

Full and Partial Retirement



in hazards suggests that individuals who fully retire are retiring at the "standard" ages close to 65 and that, conditional upon not retiring at those ages, an individual is more likely to move into partial retirement prior to full retirement than directly to full retirement.

The most important conclusion that I draw from the comparison of sample hazard rates is that the age-profiles suggest that different factors are leading to the decisions to retire fully or partially. Estimation of a model which does not distinguish between the two will confound the differences in behavior implied by the different hazards. Two cautions should be expressed at this point. First, underlying the results depicted in figure 2 is the assumption that the population is homogeneous. This assumption is obviously untenable and the econometric specifications developed in the paper are designed to relax this restriction. Second, the sample sizes become relatively small as one moves out to the right tail of the age distribution so that some caution should be taken in drawing conclusions based upon the shape of the hazards at these later ages. Still, the hazards differ by enough so that it should be safe to conclude that distinguishing between full and partial retirement is important.

2. *The Existing Literature*

Early empirical work on retirement behavior was conducted by

Boskin [1977] who analyzed the labor force participation of a 131 observation sample of households, finding very large negative effects for Social Security. Based upon estimates from a first-order Markov model of transition probabilities, he concluded that the existence of the Social Security system has increased the annual probability of retirement by 40 percent. The 40 percent figure suggests that Social Security is responsible for the bulk of the observed decline in labor force participation in the past few decades.

This view was challenged by Gordon and Blinder [1980] who, building upon their work with Wise [1980] on the institutional characteristics of Social Security, postulate a three-period life-cycle model of labor-leisure choice in which individuals presently choose to work or to retire as the market wage is greater than or less than a reservation wage. Based upon estimates from this model GS find that Social Security has rather small effects on retirement behavior--a change in the Social Security wealth to income ratio of 0.01 (about 14 percent) has a negligable impact upon retirement probabilities.[16]

Later work by, among others, Fields and Mitchell [1984], Boskin and Hurd [1984], Burkhauser [1979], Zabalza *et. al.* [1981] and

---

[16]The Gordon and Blinder finding that Social Security has almost no effect upon retirement stands virtually alone in the empirical literature, and is difficult to reconcile with the large spikes in retirement hazards coincident with Social Security eligibility found by Hausman and Wise [1985] and others.

Burtless [1986], using a variety of techniques come to a number of often conflicting conclusions. This smorgasbord of results led Aaron to conclude in 1984 that "...about all that can be said is that a preponderance of studies, whose evidentiary value is quite low, concludes that Social Security has an indeterminate impact upon retirement behavior."

Among the many studies which consider the determinants of retirement behavior, perhaps the closest to the present analysis are the recent study of partial retirement conducted by Gustman and Steinmeier [1986] and the risk models of retirement developed by Hausman and Wise [1985] and Diamond and Hausman [1984]. In this paper, I combine the concern with partial retirement of Gustman and Steinmeier with the statistical techniques developed in the Hausman and Wise and Diamond and Hausman papers, while relaxing some of the less satisfactory assumptions in both.

Gustman and Steinmeier (GS) estimate a structural model of retirement in which partial retirement is represented by the presence of an alternative wage-leisure offer. Demographic and other control variables enter into the specification solely through the preference for leisure, while economic variables enter through the lifetime budget constraint. They find moderate effects of economic variables upon retirement, with a hypothetical 50 percent increase in compensation streams for full and partial retirement, pensions, and Social Security reducing the percentage of individuals working full-time by about 10 percentage points for individuals up through age 64, and by a smaller amount for older individuals. This increase in the number of retirees is distributed unevenly across the two types of retirement.

While the GS paper is a careful analysis of the partial retirement issue and as such is an important contribution to the analysis of retirement behavior, the overly simple structure of their model makes it difficult to disentangle the effects of Social Security from the effects of, say earnings.[17] In fact, it is a requirement of the GS model that the effects of a dollar increase in the present discounted value (PDV) of Social Security benefits exactly equal a corresponding increase the PDV of earnings or wealth. If it is believed that individuals discount future Social Security benefits at a different rate than other income streams, this assumption is untenable. The assumption of identical discount rates is likely to be a problem given evidence of liquidity constraints, since individuals are prohibited by law from borrowing against Social Security wealth.[18] Furthermore, the GS assumption that individuals face a single alternative wage-leisure offer for partial retirement

---

[17]In all fairness, this is the result of the extreme complexity of their estimation technique.

[18]See Paquette [1985] for evidence that elderly individuals face constraints on their liquidity. Diamond and Hausman [1984] also provide evidence from the National Longitudinal Survey of Older Males that supports this result.

is open to criticism.

The HW and DH studies both take a different approach to analyzing the determinants of retirement. Building upon the statistical literature and the work of Lancaster [1979] they specify functional forms for the conditional probability of retirement and use these to estimate probabilities of retirement at different ages. Using a sample drawn from the RHS, HW estimate a hazard model of retirement which indicates that Social Security and other economic variables have significant effects upon retirement. According to their calculations, the increase in Social Security benefits over the past two decades may account for up to one-third of the decrease in labor force participation. DH estimate a hazard model of retirement for a slightly younger sample of National Longitudinal Survey of Older Men (NLS) individuals. Their results are similar in character to the HW results.

Unfortunately, these latter two studies assume a single form of retirement. Moreover, in the specification of the hazard functions, strong parametric restrictions are placed upon the behavior of individuals. These restrictions take the form of simple functional form specifications for the effects of aging upon retirement probabilities.[19] In this paper I relax both the assumption of a

---

[19]The same is true for the dual-risk model estimated by DH for re-employment and retirement of the unemployed. The DH restrictions involve linearity of the latent random variables for time to re-employment and time to retirement with respect to the covariates. This specification implicitly places strong, non-testable functional

single form of retirement and the parametric restrictions upon behavior.

3. *Econometric Specification*

The econometric techniques employed in this paper are derived from the extensive literature on hazard models.[20] Hazard models have enjoyed considerable popularity in empirical work since Lancaster [1979] introduced them to the economics profession in his study of unemployment durations. The previously mentioned studies by HW and DH apply these techniques to the study of retirement behavior. This paper, in contrast to earlier studies, uses techniques developed recently by Han and Hausman [1986] to estimate semi-parametric dual risk models that allow for correlation between the risks. In the retirement context, the two risks of interest are the probabilities of full and partial retirement. The technique is outlined in some detail in the remainder of the section.

It should first be noted, however, that despite an obvious relationship to other discrete choice models such as the logit, these techniques employ reduced forms which are not grounded in maximizing theory. While one can think of an individual solving a complicated

---

form restrictions upon the shapes of the baseline hazards.

[20]These models are also referred to as failure time models. Standard references in the statistical literature are Cox and Oakes [1984] and Kalbfleisch and Prentice [1980].

dynamic stochastic programming problem to determine an optimal retirement date, such a model cannot be linked directly to these hazard estimates. Intuitively, however, the model that underlies these specifications is one in which individuals solve a maximizing problem by making sequential labor-force participation decisions, comparing the utility received from retiring at a given time with the utility from continuing to work as well as the utility received from accepting an alternative, partial retirement wage-leisure offer and retiring at a later date.[21] The utility comparisons are influenced both by the age at which the decisions are being made and by the effects of other, individual specific factors.

3.1 Single Risk Estimation

Generally, an observation on a failure time can be of two types. First, the time of failure can be the direct result of the hazard process under investigation. Thus, a direct failure time observation on retirement would provide the age at which an individual is observed to retire. Alternatively, the time of failure can result from right censoring of the data. Right censoring occurs when the period over which the individual is surveyed is not long enough for

---

[21]These offers might be thought of as coming from a distribution of offers in much the same way that job offers are received by an unemployed individual (see Katz [1985]). Much more thought needs to be applied to the question of the best way in which to combine the retirement and reservation wage frameworks.

observation of a direct failure time. For example, the ending of the RHS sample period in 1979 censors failure times for those individuals who are still working at that date.

A failure time observation can therefore be characterized by the time of failure and by the type of failure, actual or censored. Suppose that failure time data of the form (t_i, δ_i, X_i), i=1,...,N, are observed where

t_i is the observed failure time for individual i

δ_i is a censoring indicator where $\delta_i = \begin{cases} 1 & \text{if a censored failure} \\ 0 & \text{otherwise} \end{cases}$

X_i is a k-vector of covariates.

In the retirement context, t represents the age of retirement if $\delta=0$, or the age that the individual leaves the survey if $\delta=1$. It should be pointed out that in most of the discussion that follows, t_i is an assumed to be a discrete approximation to the true failure time. The use of a discrete approximation results from the fact that in most data available to economists, observations on individuals are made at discrete points in time. For the RHS survey, the discreteness of the data results from the fact that observations on individual employment status may only be made at year intervals. Thus, an observed failure time of t_i implies that the true failure time lies in the interval $(t_i-1, t_i]$.

Let the continuous non-negative random variable T represent the true failure time. Suppressing the individual subscript, I assume that the conditional probability of retirement at time t can be written in the following form

$$h(t) \equiv \lim_{\Delta \to 0^+} \frac{Pr(t \leq T < t + \Delta \mid T \geq t)}{\Delta} \tag{3.1.1}$$

$$= \lambda(t)\ exp(X\beta).$$

where $\lambda(t)$ is a non-negative function of age which represents the effect of time upon the conditional probability. This model is the proportional hazards specification of Cox [1972]. It derives its name from the fact the effects of the covariates X operate proportionally upon the baseline hazard $\lambda(t)$ through the parameters β.

Two functions of $h(t)$ are of additional interest. First, the survivor function, $Q(t)$, gives the probability that the individual failure time is greater than t so that the individual is observed not to have failed at time t. In the present analysis, survival corresponds to noting that an individual has not yet retired at a given age. Second, the unconditional density function for failure time, $f(t)$, is a standard probability density function for the random variable for failure time, T. Both can readily be written in terms of $h(t)$:

$$(3.1.2) \qquad Q(t) = Pr(T \geq t) = exp(- \textstyle\int_0^t h(s)\, ds),$$

$$(3.1.3) \qquad f(t) = \lim_{\Delta \to 0^+} \frac{Pr(t \leq T < t + \Delta)}{\Delta} = h(t)\, Q(t).$$

The complete specification of the likelihood of an observation for either continuous or discrete data is possible using a combination of (3.1.2) and (3.1.3).

Unobservable individual effects are readily incorporated into the specification. Suppose that there is an additional random component to the hazard function, θ, which enters multiplicatively. Then (3.1.1) becomes

$$(3.1.4) \qquad h(t|\theta) = \theta\, \lambda(t)\, exp(X\beta) = \theta\, h(t).$$

where θ is a non-negative random variable assumed to be independent of the covariates. Following the technique of Lancaster [1979], assuming that θ is distributed as a unit-mean gamma random variable and taking expectations with respect to θ allows for the derivation of closed-form expressions for the survivor function

$$\begin{aligned}(3.1.5) \qquad Q^*(t) &\equiv E_\theta\, Q(t|\theta) \\ &= E_\theta\, exp(- \textstyle\int_0^t h(s|\theta)\, ds), \\ &= (1 + \sigma^2 \textstyle\int_0^t h(s)\, ds)^{-1/\sigma^2},\end{aligned}$$

where σ^2 is the variance of the gamma distribution.

It is a straightforward exercise to write out, for the entire sample, the log-likelihood functions associated with each of these specifications. For illustrative purposes, I consider two cases. First, if there is continuous data and no heterogeneity, the log-likelihood function is given by

$$(3.1.6) \quad \log L(\beta) = \sum_{i=1}^{N} (1-\delta_i) \log h(t_i) + \delta_i \log Q(t_i) \ ,$$

while if heterogeneity is present and there is discrete data of the form described above, the log-likelihood is

$$(3.1.7) \quad \log L(\beta,\sigma) = \sum_{i=1}^{N} (1-\delta_i) \log (Q^*(t_i-1) - Q^*(t_i)) + \delta_i \log Q^*(t_i) \ .$$

Most commonly, the specification has been completed by choosing a functional parameterization of the baseline hazard $\lambda(t)$--often a single-parameter Weibull ($\lambda(t) = \alpha\ t^{\alpha-1}$). Unfortunately, it has been established that results obtained from maximum likelihood estimation of these functions are sensitive to the choice of a distribution for the heterogeneity parameter and the choice of functional form for the baseline hazard.[22] To reduce the biases resulting from functional form restrictions, I adopt the semi-parametric HH approach of estimating (3.1.6) and (3.1.7) parametrically with respect to the β, but with no restrictions on the form of $\lambda(t)$.

To see how the technique works, note that (3.1.3) can be rewritten as the transformed model

$$l_t = -X\beta + \epsilon, \tag{3.1.8}$$

$$l_t \equiv log \int_0^t \lambda(s)\, ds \tag{3.1.9}$$

where ϵ has an extreme value distribution. Rewriting the likelihood in the regression form involves a simple transformation of variables, a process which is outlined in greater detail in Appendix 1.

Suppose now that an individual is observed to have retired at time t where t is again a discrete approximation to the actual failure time T. The probability of observing a failure at time t for a given individual is given by

$$Pr(t-1 \leq T < t) = \int_{l_{t-1}+X\beta}^{l_t+X\beta} f(\epsilon)\, d\epsilon \quad . \tag{3.1.10}$$

The estimation technique involves treating the l_t-functions of the

---

[22] See Heckman and Singer [1985] and Manton, *et. al*. [1986].

baseline hazard as a parameter for each of the T potential failure periods and estimating them jointly with the β. Letting $y_{it} = 1$ if a failure is observed for individual i at time period t and 0 otherwise, the log-likelihood for the entire sample corresponding to (3.1.6) can be written as

$$(3.1.11) \quad \log L(\beta, l) = \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it} \left(\delta_i \log \int_{l_t + X\beta}^{\infty} f(\epsilon)\, d\epsilon + (1-\delta_i) \log \int_{l_{t-1}+X\beta}^{l_t + X\beta} f(\epsilon)\, d\epsilon \right).$$

where $(l = \{l_1, \ldots, l_T\})$. Since ϵ has an extreme value distribution, the specification corresponds to an ordered logit likelihood, with the categories determined by the T failure times.[23] The addition of heterogeneity to the model is straightforward, and is presented in Appendix 2.

3.2 *Competing Risks Estimation*

The extension of the semi-parametric technique to the competing

---

[23]As might be expected, assuming that the ϵ has a standard normal distribution yields essentially the same parameter estimates (after scaling for the unequal variances) except in the extreme tails of the distributions. Both models are easy to compute since algorithms for computing the *CDFs* are well-known. This finding is analagous to the results for the logit and probit models.

risk model is natural. Suppose that the model is cast in terms of latent variables. Let $k = 1,\ldots,K$ represent the K competing risks in the model. Let the non-negative random variables Y_k^* represent the possibly latent failure times for each of the K risks. These random variables are analagous to the random variable T defined above for the single risk case. Then let the observed failure time T be the minimum of the underlying durations so that

$$(3.2.1) \qquad T = min(Y_1^*,\ldots,Y_K^*)$$

represents the observed failure time. Thus, in a manner parallel to the single risk case, cause-specific hazard rates can be specified

$$(3.2.2) \qquad h_k(t) \equiv \lim_{\Delta\to 0^+} \frac{Pr(t \le Y_k < t + \Delta \mid Y_j \ge t,\ j=1,\ldots,K)}{\Delta},$$

$$= \lambda_k(t)\ exp(X_k\beta_k),$$

$k = 1,\ldots,K$. In the presence of all K risks and no ties, the overall hazard, survivor and density functions are given by

$$(3.2.3) \qquad h(t) = \sum_{k=1}^{K} h_k(t),$$

$$(3.2.4) \qquad Q(t) = exp(- \textstyle\int_0^t h(s)\ ds) ,$$

$$f_k(t) = h_k(t)\ Q(t)\ . \tag{3.2.5}$$

$k=1,...,K$.

Once again, these functions can be expressed in their regression forms. For simplicity, consider the case of two risks. Adopting the notation used for the single risk specification, the model can be written as

$$I_t^1 = -X_1\beta + \epsilon_1, \tag{3.2.6}$$

$$I_t^2 = -X_2\beta + \epsilon_2.$$

If ϵ_1 and ϵ_2 are assumed to be independent, the likelihood function for each risk can be estimated as a single risk model in which the alternative failures are treated as censored observations.[24] If, for example, the errors are assumed to be independent Weibull, then the techniques described in the previous section are directly applicable.

If the errors are not independent, then the joint distribution of (ϵ_1,ϵ_2) is required. The probability of observing a failure of

[24]This technique is similar to the approach taken by Katz [1986] who estimates models based upon the Lancaster-type likelihood, taking account of factoring of the joing likelihood under the assumption of independence. He estimates his model assuming the baseline hazard is a modified Weibull.

type 1 at $t = min(t_1, t_2)$ is the probability that t_1 lies in the interval $(t\text{-}1, t]$ and t_2 is greater than t_1 as the latter is evaluated at every point in that interval. The probability is represented by

$$(3.2.7) \qquad H_1(t) = \int_{I^1_{t-1}+X\beta}^{I^1_t \;\; +X\beta} \int_{g(\epsilon_1)}^{\infty} f(\epsilon_1, \epsilon_2)\, d\epsilon_2\, d\epsilon_1$$

where $g(\epsilon_1)$ is the function relating the two errors so that t_2 is equal to t_1. An analagous expression results if the failure is of type 2. Assuming linearity of the hazards over the interval, it can be shown that

$$(3.2.8) \qquad g(\epsilon_1) = (I^2_t + X_2\beta_2) + \frac{(I^2_t - I^2_{t-1})}{(I^1_t - I^1_{t-1})} \; (\epsilon_1 - (I^1_t + X_1\beta)),$$

(Appendix 3).

If an observation is censored so that there is no failure of either type, the corresponding likelihood is given by the joint survivor function

$$(3.2.9) \qquad C(t) = \int_{I^1_{t-1}+X_1\beta_1}^{\infty} \int_{I^2_{t-1}+X_2\beta_2}^{\infty} f(\epsilon_1, \epsilon_2)\, d\epsilon_2\, d\epsilon_1.$$

Completion of the specification is accomplished by the choice of $f(\epsilon_1, \epsilon_2)$ to be bivariate normal with correlation coefficient ρ. The specification allows for dependence between the disturbances in the model. Since the baseline hazards for the two risks are estimated semi-parametrically, the strong functional form restrictions made by DH and HW are avoided.[25]

The observed data are now of the form $(t_i, s_i(k), \delta_i)$, k=1,2, where t and δ are as defined before and s is an indicator variable for the type of failure ($s(k)=1$ if failure of type k, $s(k)=0$ otherwise). The indicator variable y_{it} is defined as in the single risk case. The log-likelihood function can be written as

$$(3.2.10)\quad \log L(\beta, \lambda, \rho) = \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it} \left(\delta_i \log C(t) + \sum_{k=1}^{2} s_i(k) \log H_k(t)\right).$$

The likelihood function is therefore maximized over the parameters β_1, β_2, the semi-parametric baseline hazards ($\lambda^k = \{\lambda_1^k, \ldots, \lambda_T^k\}$, k=1,2) and the correlation coefficient for the bivariate normal ρ.

All of the likelihoods in this paper were maximized using standard maximum likelihood techniques. The algorithm used is a modified version of the Berndt, Hall, Hall and Hausman [1974]

---

[25]The extension to several risks is straightforward. With present computational techniques, estimation of models with up to three or four risks should be possible. See Hausman and Wise [1978] for a discussion of the computational issues relating to the evaluation of multivariate normal distributions. See also McFadden [1986] for a simulated moments approach to the estimation of models with many risks.

gradient technique that employs quadratic interpolation to determine stepsizes. Convergence for all of the models was achieved with little difficulty. The actual estimation technique involved estimating the β and l with retirement ages past a given age treated as censored failures. The estimated parameters were then used as starting values for a model with all retirement ages unconstrained.

4. Data

The data used in this study are derived from the Longitudinal Retirement History (RHS). The RHS is a panel data set comprised of observations at two-year intervals on 11,000 elderly individuals, aged 58-63 in 1969 over the period from 1969-1979. The sample used in the estimation is restricted to male, non-farm workers in private, but not self-employment, for whom there are sufficiently complete data to allow for observation over at least one of the five RHS waves (1971-1979) that are used in this study. The original extract that I received was comprised of approximately 4,000 observations; after considerable cleaning, the data set used in the estimation consists of 1,633 individuals.

The first wave of the RHS (1969) was not used because of problems with the comparability of data across the surveys. Of the original 4,400 observations, 697 were lost because of an inability to determine industrial and occupation classifications. 384 were lost because of insufficient data to impute assets holdings for any of the five waves. There were 199 individuals who had highly irregular work histories. An additional 1,474 individuals were lost because I am unable to follow their employment status for at least one period or because of poor data over the sample period. For example, for 138 individuals it was impossible to use existing earnings records or to impute earnings records.

In principle, one could use the information that individual had not retired prior to the date that the poor data were observed--this would involve censoring the individual failure time at that date. This would, however, have the effect of mixing the "failure times" of the poor data with the hazard processes for full and partial retirement. Moreover, given the computational complexity of the dual risk likelihood and the resulting desire to keep sample sizes small, I chose to leave that extension for future research. The following results should be interpreted bearing this sample selection proceedure in mind.

Since the nature of work effort is central to the analysis, it is important to distinguish carefully between full and partial retirement. As noted above, an hours/weeks based definition of the extent of work is used. Individuals who are reported as fully or partially retiring in a given sample year are traced back through the previous sample period to see whether or not they have held an intervening partial retirement job, with the type of retirement and age at retirement adjusted accordingly. Ages at retirement are then incremented by one so that a reported retirement age of 65 indicates that the individual retired in the interval (64,65]. The age adjustment is merely a normalization and is for computational convenience.

Two types of variables are used in the estimation procedure: heterogeneity controls and economic variables. The control variables include health status, number of persons in household, marital status, race, education, the presence of mandatory retirement provisions on the job, and occupation class. The economic variables are related to the income and compensation streams that individuals receive and include earnings, wealth, pension eligibility and Social Security.

Most of the control variables are self-explanatory. Health status is an indicator for poor health based upon self-reported status. The responses are based upon the answer to the question, "Is your health worse than others your age?" It has long been thought that health status is one of the primary determinants of retirement age. The effect of changes in health status will be to change preferences for leisure and perhaps to create an exogenous change in the wage rate. Poor health should therefore have a positive impact upon retirement probabilities. Additionally, it might be expected that health status will have a different effect upon the probabilities of full and partial retirement since individuals with poor health should be more likely to retire fully than partially retire given that the distribution of partial retirement wage-leisure offers should be poorer for an individual with bad health. In contrast, the characteristics of the full retirement choice are unchanged.

It should be noted that a number of researchers have questioned the exogeneity of self-reported health measures, arguing that individuals will report poor health to avoid the stigma associated with early retirement. The issue of health status endogeneity is raised by Parnes [1981] and discussed at greater length by Feldman [1984]. The latter concludes that the reporting error is likely to be of little practical importance. Furthermore, DH find that the use of different indicators of health status does not alter their results appreciably.

The economic variables follow closely the variables used in existing studies. Earnings are included as a measure of the opportunity cost of retiring. For individuals with adequate information, actual earnings are used, otherwise earnings are imputed on the basis of the Social Security earnings records that have been adjusted by the Fox [1981] imputation method to account for truncation. These values are net of payroll taxes but not of income taxes because of serious difficulties with measuring non-labor income.[26] Theory suggests that earnings will reduce the probability

---

[26]This use of gross earnings levels is not unique despite the conflict with theory. The addition of properly computed income taxes is difficult given the complexity of the RHS, but is nevertheless a

of retiring since high earnings represent a high cost of retiring. The assets data are based upon estimates of the non-housing wealth held by individuals. Since the RHS data is somewhat spotty, where it is necessary these values are imputed based upon asset shares equations estimated for those who have complete data.[27] It is expected that assets will have a positive impact upon retirement probabilities for the reasons noted previously.

Private pensions are entered via indicators for eligibility. Eligibility indicators for both full and reduced pension benefits are considered with both expected to increase the probability of retirement. Ideally, the dollar amounts of these pensions would be available, but the RHS data on pensions would require considerable cleaning to support such calculations. I decided that the eligibility indicators were preferable to additional imputation based upon general pension tables or estimated values.

Social Security benefits are entered in two ways. First, a measure of the benefits that the individual would receive if he retired at age 62 is calculated. In computing these benefits, the Social Security rules for the given year and cohort are applied to the individual earnings records. First, AMWs are calculated based

priority for future research.

[27]These data, as well as the original RHS extract, were generously provided by Beverly Hirtle.

upon earnings histories, then the appropriate PIA is calculated. One advantage to using the RHS is that it contains matched earnings records for individuals from 1951-1974 so that benefit levels can, for the most part, be calculated on the basis of actual earnings. If appropriate, PIAs are also computed for spouses based upon their earnings histories. I then calculate benefit levels using the appropriate rules for one or two-earner housholds. Note that it is very important to match properly the age and cohort with the contemporaneous law since the period of observation is one of greatly varying Social Security rules and benefit levels. I approximate the effects of ABR by calculating benefit levels were the individual to retire at age 65, and then looking at the difference between the two levels. The corresponding "delta" variables are entered into the specification along with the age 62 benefit levels.[28]

Descriptive statistics for the control and economic variables are presented below in table 5.

5. *Results*

In this section, I present results for several models of

---

[28]The use of age 65 for the delta calculation results from the institutional importance of ages 62 and 65 for the computation of Social Security benefits. Delta variables are also calculated for ages 64 and 66. While these variables were not used in the estimation procedure, the high correlation with the age-65 delta (0.79 and 0.71) suggests that the results should not change much using alternative measures.

Table 5 - Variable definitions and descriptive statistics, RHS subsample.

Variable	Description	Mean (Std. Dev.)
Control Variables		
General Controls		
education	= years of education	10.805 (3.447)
health	= 1 if health worse than cohort	0.319 (0.466)
married	= 1 if married with spouse present	0.890 (0.313)
non-white	= 1 if non-white	0.097 (0.296)
household size	= number of individuals in household	2.532 (1.240)
mandatory retire	= 1 if mandatory retirement provisions present	0.372 (0.484)
Occupation Controls		
clerical	= 1 if clerical, service, sales	0.219 (0.413)
craft	= 1 if craftsman	0.247 (0.431)
labor	= 1 if laborer, operative	0.292 (0.455)
manager	= 1 if manager	0.129 (0.335)
professional	= 1 if professional, technical	0.067 (0.251)

Table 5 - (continued)

Economic Variables		
earnings	= earnings (x $1,000)	6.926 (4.800)
wealth	= non-housing assets (x $10,000)	1.025 (3.194)
full pension	= 1 if full pension eligibility	0.141 (0.348)
part pension	= 1 if partial pension eligibility	0.160 (0.367)
ss62	= monthly social security benefits if retire at age 62 (x $100)	2.295 (0.644)
delta	= change in social security benefits from ss62 if retire at 65 (x $10)	2.004 (2.627)

Note: Based upon author's calculations. Amounts are expressed in real 1967 dollars.

retirement. While the principle contribution of this paper is the estimation of a competing risks model of retirement which allows for correlation between risks, for purposes of comparison I estimate single risk hazards for both the parametric and semi-parametric models. The single risk models correspond to considering a single-form of retirement in which an individual is classified as retired if he either fully or partially retires. I then discuss differences between the estimates derived from the single risk parametric, and semi-parametric models. Two competing risk models of retirement are considered: first, one in which the risks are assumed to be uncorrelated, and second, a model where I allow for correlation between the stochastic disturbances.

5.1 Single Form of Retirement

5.1.1 Parametric Baseline Hazard (Weibull)

For purposes of comparison with the competing risk case, I calculate parameter estimates for single risk, parametric models of retirement in both the no heterogeneity and heterogeneity cases. The baseline hazard for the single retirement model is assumed to be simple Weibull ($\lambda(t) = \alpha\ t^{\alpha-1}$). The results are presented in table 6.

Interpretation of the coefficients in table 6 is aided by the fact that $log\ t^{\alpha}$ has an extreme value distribution so that t^{α} is

Table 6 - Parameter estimates for single form of retirement model, weibull baseline hazard.

Variable	No Heterogeneity (1)	(2)	(3)	Heterogeneity (4)	(5)	(6)
Control Variables						
constant	-6.780	-7.381	-7.873	-9.584	-10.518	-10.790
	(0.187)	(0.208)	(0.223)	(0.330)	(0.341)	(0.366)
education	-0.035	-0.046	-0.026	-0.042	-0.054	-0.034
	(0.008)	(0.008)	(0.009)	(0.014)	(0.015)	(0.016)
health	0.339	0.353	0.334	0.463	0.495	0.483
	(0.053)	(0.053)	(0.053)	(0.095)	(0.098)	(0.095)
married	0.109	-0.344	-0.319	-0.073	-0.582	-0.533
	(0.073)	(0.088)	(0.092)	(0.127)	(0.156)	(0.153)
non-white	-0.231	-0.159	-0.159	-0.248	-0.164	-0.164
	(0.084)	(0.088)	(0.090)	(0.146)	(0.152)	(0.151)
household size	0.024	0.023	0.025	0.059	0.071	0.062
	(0.020)	(0.020)	(0.021)	(0.034)	(0.066)	(0.064)
mandatory retire	0.525	0.435	0.396	0.542	0.411	0.382
	(0.052)	(0.052)	(0.054)	(0.094)	(0.098)	(0.097)
clerical	-----	-----	0.024	-----	-----	0.074
			(0.122)			(0.200)
craft	-----	-----	0.364	-----	-----	0.336
			(0.119)			(0.202)
labor	-----	-----	0.393	-----	-----	0.436
			(0.114)			(0.195)
manager	-----	-----	-0.113	-----	-----	-0.184
			(0.134)			(0.229)
professional	-----	-----	-0.070	-----	-----	-0.093
			(0.146)			(0.253)

Table 6 - (continued)

Variable	No Heterogeneity (1)	(2)	(3)	Heterogeneity (4)	(5)	(6)
Economic Variables						
earnings	0.006	-0.013	-0.007	0.006	-0.017	-0.009
	(0.004)	(0.005)	(0.006)	(0.009)	(0.010)	(0.011)
wealth	0.002	0.015	0.022	-0.001	0.023	0.031
	(0.012)	(0.013)	(0.014)	(0.019)	(0.017)	(0.019)
full pension	0.073	0.065	0.046	0.066	0.063	0.072
	(0.071)	(0.072)	(0.071)	(0.123)	(0.127)	(0.123)
part pension	0.015	0.006	0.033	0.099	0.055	0.044
	(0.059)	(0.062)	(0.067)	(0.113)	(0.117)	(0.115)
ss62	-----	0.488	0.450	-----	0.635	0.588
		(0.051)	(0.053)		(0.091)	(0.091)
delta (ss65-ss62)	-----	-0.019	-0.023	-----	-0.077	-0.075
		(0.010)	(0.010)		(0.018)	(0.018)
Other						
σ	-----	-----	-----	0.905	0.954	0.914
				(0.045)	(0.044)	(0.046)
alpha	3.097	3.190	3.251	4.752	5.040	4.952
	(0.061)	(0.064)	(0.065)	(0.152)	(0.155)	(0.155)
log likelihood	-3689.19	-3648.42	-3622.58	-3603.82	-3562.64	-3550.29

Note: Asymptotic standard errors in parentheses. Retirement is defined to be the first occurence of either full or partial retirement. The expected value of the duration $E(t)$ is given by the expression $E(t) = \Gamma(1 + 1/\alpha)\ exp(-X\beta/\alpha)$ where Γ is the incomplete gamma function. This implies that $dln(E(t))/dX = -\beta/\alpha$ so that a unit change in the independent variables yields a percentage change in the expected duration of $-\beta/\alpha$. Note that these durations are normalized with age 58 equal to 1.

distributed as a standardized exponential. It follows that $E(t) = \Gamma(1/\alpha + 1)\ exp(-X\beta/\alpha)$ where Γ is the incomplete gamma function, so that $log\ E(t) = log\ \rho(1/\alpha + 1) - X\beta/\alpha$. Then a unit change in X results in a percentage change in the mean duration of $-\beta/\alpha$. Note that these percentage changes are calculated with age 58 normalized to be 1.

In general, the results are similar to those obtained in previous studies. The economic variables have, for the most part, the expected signs. Both Social Security and the Social Security delta have strong, precisely measured effects upon retirement. A $25 increase in monthly Social Security benefits reduces the normalized mean retirement age by about 13 percent in specification (3) ($-\beta/\alpha$ = 0.450/3.215). Since the mean of the normalized age is about 8, a rough calculation indicates that this corresponds to a change in the mean age of a little over a year (dt/t = .13 ; dt = .13 x 8). In contrast, an identical increase in the Social Security delta has about half that effect in raising the average duration. Once Social Security benefits are accounted for, additional earnings reduce slightly the probability of retiring, in specification (2) raising $E(t)$ by about 2 months for every $2,500 increase. It is interesting to note that the impact of earnings drops by about a half with the addition of occupation controls, suggesting that it is difficult to disentangle earnings effects from those resulting from job class. Additional wealth and pension eligibility all increase the retirement hazards, but these parameters are estimated with considerable error. It may be that the RHS data are not sufficiently rich to support strong inferences about the effects of these variables, or that the effects are relatively small.

Consider next the influence of the control variables. Health status is one of the more important of these variables. Focusing for a moment on specification (2), bad health reduces the mean age of retirement by approximately 10 percent ($-\beta/\alpha$ = 0.353/3.190), or about 11 months. Similarly, the presence of mandatory retirement provisions on the primary job strongly increases retirement probabilities. In contrast, additional education, being married and non-white all increase the mean retirement age. The coefficient on education, implying about a 1 month increase in the mean age per year of additional schooling, is small relative to the findings of HW and DH, but the magnitude may result from the attempt to estimate separately the impact of race and education. The occupation control variables in specification (3) suggest that being a laborer or craft worker strongly increase the probability of retirement, with the effect comparable to that resulting from bad health. Other occupation variables suggest more moderate effects, with managers and professionals experiencing a lower retirement hazard, and clerical workers a higher hazard. The latter occupation parameters are, however, measured imprecisely. These results are relatively consistent across specifications.

5.1.2 Semi-Parametric Baseline Hazard

The results of estimation of the semi-parametric likelihood for the single form of retirement model are given in table 7 for the no-heterogeneity specification. I consider this model to see whether estimating a semi-parametric baseline hazard yields results that differ appreciably from those given by the Weibull specification.

The parameter estimates for the β in table 7 are similar to those for the Weibull-parametric hazard presented in table 6. If there is a pattern to the differences, it is that the effects of independent variables are almost always smaller in the semi-parametric baseline hazard model than in the Weibull model. For example, comparing specification (3) in tables 6 and 7, the effects of Social Security benefits are about 10 percent lower in the semi-parametric case; the parameter for health status is also about 10 percent lower. Similar differences are found for the effects of most of the variables, the exceptions being those for the imprecisely measured pension benefits. The result that estimating the baseline hazard semi-parametrically lowers parameter estimates is intuitive, since allowing more variation in the estimated baseline hazard will reduce the apparent influence of covariates.

The estimated baseline hazard rate for the semi-parametric model is compared with the corresponding Weibull hazard rate in figure 3. While similar up through age 65, the estimates differ markedly in the

Table 7 - Parameter estimates for single form of retirement model, semi-parametric baseline hazard without heterogeneity controls.

Variable	(1)	(2)	(3)
Control Variables			
education	-0.031	-0.041	-0.024
	(0.009)	(0.009)	(0.011)
health	0.310	0.327	0.318
	(0.059)	(0.059)	(0.059)
married	0.090	-0.308	-0.287
	(0.085)	(0.101)	(0.103)
non-white	-0.202	-0.143	-0.141
	(0.095)	(0.098)	(0.099)
household size	0.015	0.015	0.015
	(0.022)	(0.022)	(0.022)
mandatory retire	0.475	0.398	0.367
	(0.058)	(0.057)	(0.058)
clerical	-----	-----	0.023
			(0.141)
craft	-----	-----	0.284
			(0.140)
labor	-----	-----	0.312
			(0.136)
manager	-----	-----	-0.115
			(0.157)
professional	-----	-----	-0.054
			(0.173)

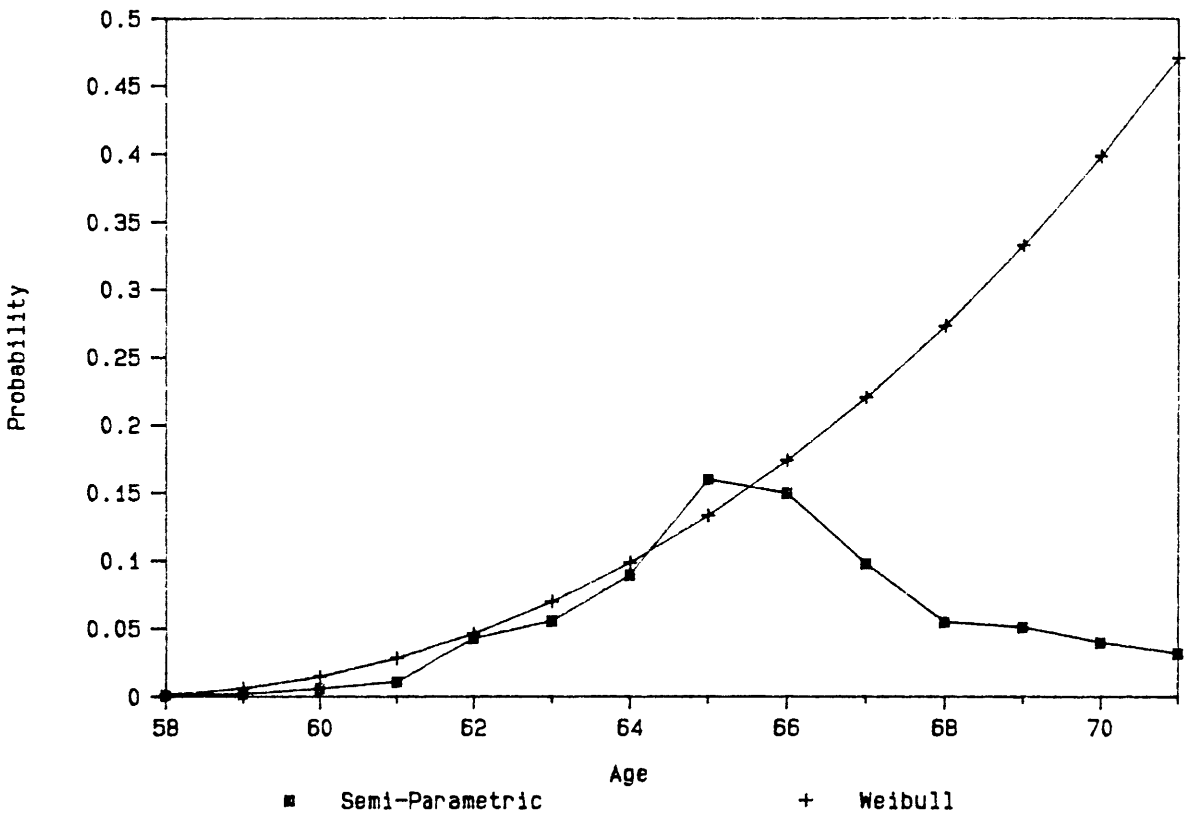
Table 7 - (continued)

Variable	(1)	(2)	(3)
Economic Variables			
earnings	0.004	-0.013	-0.009
	(0.005)	(0.007)	(0.007)
wealth	0.001	0.013	0.020
	(0.013)	(0.014)	(0.014)
full pension	0.071	0.062	0.053
	(0.081)	(0.081)	(0.080)
part pension	0.041	0.031	0.050
	(0.072)	(0.073)	(0.075)
ss62	-----	0.433	0.407
		(0.057)	(0.059)
delta (ss65-ss62)	-----	-0.018	-0.022
		(0.012)	(0.012)
log likelihood	-3501.65	-3468.95	-3452.11

Note: Asymptotic standard errors in parentheses.

Figure 3 - Estimated Hazard Rates for Retirement

Weibull vs. Semi-Parametric



post age-65 period. The Weibull, because of the simple parametric form, shows the conditional probability of retirement to be increasing for individuals past age 65. In contrast, the semi-parametric model has a decreasing hazard over the same period. The difference in post age-65 hazard rates results from the fact that there are relatively fewer individuals retiring at those ages so that the maximization algorithm for the Weibull places more weight upon fitting the increasing part of the hazard. The striking difference in the estimated baseline hazards implies that inferences about the effects of variables upon retirement behavior will be affected by the shapes of the hazards as well as differences in the coefficients between the two specifications.

Given the semi-parametric estimates of the baseline hazard, the assumption of a Weibull hazard rate can be tested using standard minimum χ^2 techniques. Denoting the semiparametric estimates as τ and the corresponding discrete points in the true baseline hazard as λ_o, standard maximum likelihood theorems show that, asymptotically, $\tau \sim N(\lambda_o, \Omega)$, where Ω is the appropriate block of the inverse of the Fisher information matrix. I fit a two-parameter Weibull to this model, $g(\alpha,\gamma) = \gamma\ t^{\alpha-1}$. Then the quadratic form

$$W = (\tau - g(\hat{\alpha},\hat{\gamma}))'\ \Omega^{-1}\ (\tau - g(\hat{\alpha},\hat{\gamma}))$$

is distributed under the null hypothesis of the Weibull specification as asymptotic χ^2 with k-2 degrees of freedom where k is the number of parameters in the estimated baseline hazard. The value of W is estimated to be 166, which is well in excess of a 5% critical value for the χ^2 so that the Weibull assumption is rejected.[29]

To see the importance of the difference in hazard rates implied by the two models, I perform a crude simulation of the effects of a large change in the Social Security system. Starting with the sample of 381 individuals who were 58 years old in 1969, I first calculate the sample survivor function for those individuals given their actual Social Security benefits, calculated under the rules that were in effect for that cohort.[30] For comparison, I then calculate the survivor function under the relevant Social Security law for 1969. The simulation yields an approximate measure of the extent of labor force participation were the system to have remained unchanged. The differences between the two estimates of the survivor function and the corresponding hazards provide a rough measure of the impact of the changes in Social Security law in the early 1970s.

The simulated survivor functions for the Weibull and semi-parametric estimates are presented in table 8. For both models,

---

[29] A natural alternative test is to perform a Hausman-type specification test of the β (Hausman [1978]). Under the null hypothesis of a Weibull specification, the parametric model is both consistent and efficient for the β, while the semi-parametric is consistent under the null and the alternative, but not efficient.

[30] In the remainder of the discussion, these rules will be referred to as those existing under "current law".

Table 8 - Estimated sample survivor functions under current and 1969 Social Security law, various models. Averages over 381 individuals

Single Risk Models

	Weibull		Extreme Value	
Age	Current Law	1969 Law	Current Law	1969 Law
58	0.9987	0.9989	0.9961	0.9969
59	0.9873	0.9900	0.9915	0.9931
60	0.9535	0.9631	0.9731	0.9781
61	0.8865	0.9091	0.9416	0.9523
62	0.7821	0.8226	0.8300	0.8592
63	0.6463	0.7057	0.6969	0.7444
64	0.4956	0.5685	0.5221	0.5859
65	0.3508	0.4271	0.2942	0.3612
66	0.2289	0.2982	0.1568	0.2101
67	0.1381	0.1933	0.0963	0.1372
68	0.0774	0.1165	0.0700	0.1034
69	0.0404	0.0654	0.0504	0.0771
70	0.0197	0.0343	0.0374	0.0590
71	0.0090	0.0168	0.0287	0.0466

Independent Competing Risks Model

	Full Retirement		Partial Retirement	
Age	Current Law	1969 Law	Current Law	1969 Law
58	0.9974	0.9982	0.9993	0.9991
59	0.9937	0.9956	0.9906	0.9881
60	0.9862	0.9902	0.9795	0.9740
61	0.9710	0.9791	0.9302	0.9120
62	0.9248	0.9451	0.8815	0.8518
63	0.8571	0.8946	0.7577	0.7029
64	0.7905	0.8433	0.6796	0.6124
65	0.6056	0.6904	0.5839	0.5055
66	0.4661	0.5629	0.4763	0.3910
67	0.4001	0.4982	0.4247	0.3384
68	0.3550	0.4520	0.3647	0.2795
69	0.3209	0.4159	0.3363	0.2525
70	0.2826	0.3742	0.2851	0.2053
71	0.2652	0.3547	0.2728	0.1942

Table 8 - (continued)

Correlated Competing Risks Model

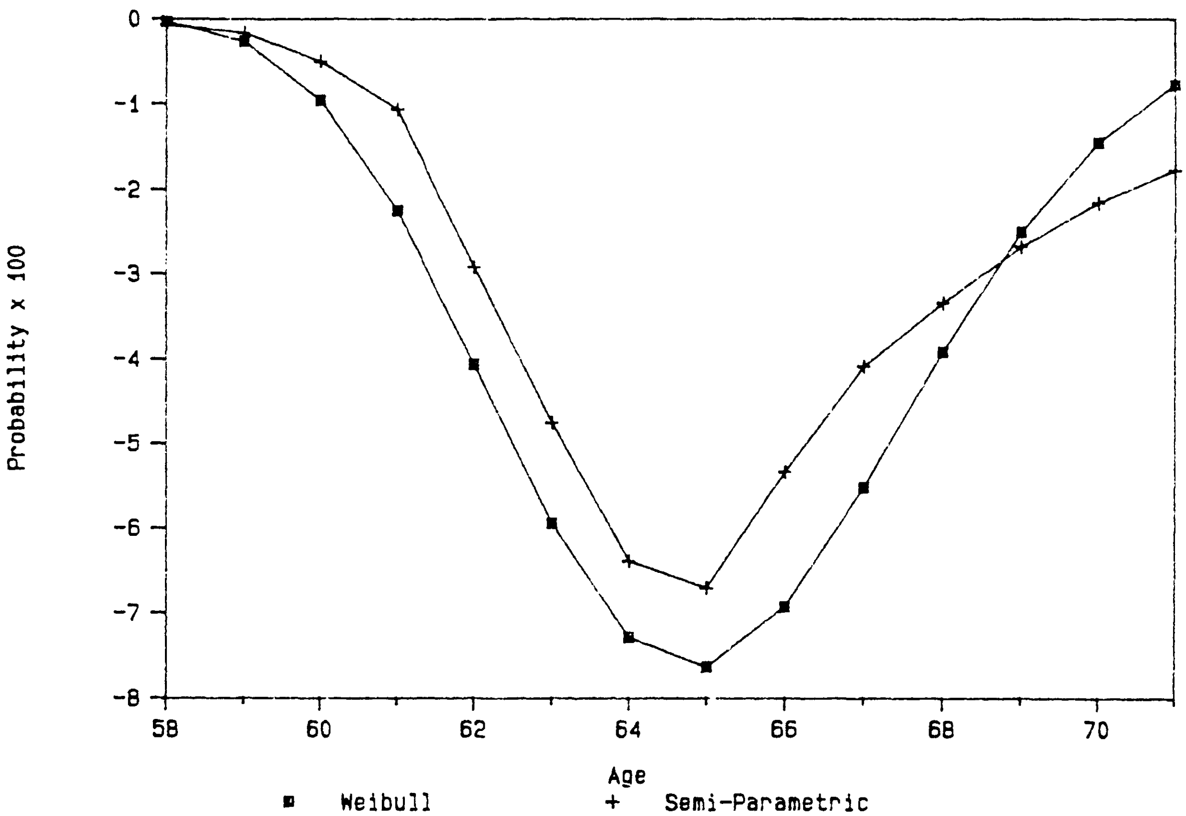
Age	Full Retirement Current Law	Full Retirement 1969 Law	Partial Retirement Current Law	Partial Retirement 1969 Law
58	0.9941	0.9964	0.9980	0.9978
59	0.9869	0.9916	0.9980	0.9978
60	0.9719	0.9811	0.9796	0.9777
61	0.9401	0.9574	0.9594	0.9560
62	0.8420	0.8780	0.8678	0.8592
63	0.7113	0.7632	0.7762	0.7640
64	0.5854	0.6454	0.5779	0.5613
65	0.3191	0.3751	0.4485	0.4313
66	0.1828	0.2245	0.2880	0.2724
67	0.1273	0.1601	0.1782	0.1656
68	0.0952	0.1218	0.1405	0.1293
69	0.0741	0.0960	0.1046	0.0951
70	0.0548	0.0720	0.0895	0.0809
71	0.0464	0.0615	0.0657	0.0586

Note: The sample survivor functions are calculated by averaging the survivor functions for individuals across the 381 individuals in the youngest cohort of the sample. The calculation for 1969 law involves simulating benefits for individuals over time, given no change from the 1969 rules, and then computing the implied survival probabilities.

the change in Social Security law has moderate effects upon survival probabilities, with changes in survivor rates mostly in the range of 4-7 percent. Of the two specifications, the Weibull model generally yields larger changes in the survivor function. These changes in the survivor function are presented graphically in figure 4. The results indicate that the changes in the Social Security system have had moderate effects upon retirement probabilities. The Weibull model simulation yields survivor probabilities for a 62 year old individual that are about 4 percentage points higher under 1969 law than under the actual law (82.26-78.21) implying higher participation rates under the earlier law. The gap between the Weibull survivor functions for the two laws peaks at age 65 at around 7.5 percent, then decreases smoothly to less than 1 percent at age 71.

As noted above, despite a similar age-profile, the changes in survivor functions associated with the semi-parametric model are generally smaller than those for the Weibull model. The differences result directly from the smaller parameter estimates for the Social Security and the delta variables in the semi-parametric model. The differences between the two simulations are sizeable since for the ages 61-67, the response of the Weibull survivor is almost uniformly one percentage point higher. Thus, the semi-parametric model generates smaller changes in participation rates resulting from changes in the Social Security system. Once again, the fact that the Weibull model yields larger simulated changes in the survivor

Figure 4 - Change in Survivor Function Resulting From Change in Social Security Law, Single Form of Retirement Models



function is not surprising. Estimating a proportional hazards model with a parametric restriction on the baseline hazard means that some of the variation over time in the baseline hazard that is not captured by the parametric specification will instead be attributed to the effects of covariates.

Associated with the survivor functions depicted above are simulated values for the sample hazard rates. These hazard rates are given in table 9. Changes in the hazard rates resulting from changes in the Social Security rules are depicted in figure 5. The simulation predicts moderate changes in the conditional probabilities of retirement resulting from the change from 1969 Social Security rules--for both models, in the 62-66 age range, the hazard rate is from 2 to 5 percentage points higher than under 1969 law. The smooth shape of the difference in the Weibull hazards is in contrast to the corresponding difference in the semi-parametric hazards which displays a sharply defined spike at age 65. In addition, the semi-parametric model exhibits slightly larger sensitivity to the change in Social Security law. For example, at age 65, the change from 1969 Social Security rules to the actual rules yields in excess of a 5 percentage point increase in the hazard rate for the semi-parametric model, and a 4 percentage point increase under the Weibull model.

Table 9 - Estimated sample hazard rates under current and 1969 Social Security law, various models. Averages over 381 individuals.

Single Risk Models

Age	Weibull Current Law	Weibull 1969 Law	Extreme Value Current Law	Extreme Value 1969 Law
58	0.0013	0.0011	0.0039	0.0031
59	0.0114	0.0090	0.0046	0.0038
60	0.0114	0.0090	0.0186	0.0151
61	0.0342	0.0271	0.0323	0.0264
62	0.0703	0.0561	0.1186	0.0978
63	0.1178	0.0951	0.1603	0.1336
64	0.1736	0.1421	0.2508	0.2129
65	0.2332	0.1944	0.4366	0.3835
66	0.2923	0.2487	0.4670	0.4183
67	0.3474	0.3019	0.3855	0.3471
68	0.3966	0.3518	0.2738	0.2466
69	0.4397	0.3972	0.2800	0.2539
70	0.4776	0.4383	0.2582	0.2352
71	0.5119	0.4760	0.2306	0.2107

Independent Competing Risks Model

Age	Full Retirement Current Law	Full Retirement 1969 Law	Partial Retirement Current Law	Partial Retirement 1969 Law
58	0.0026	0.0018	0.0007	0.0009
59	0.0037	0.0026	0.0086	0.0110
60	0.0076	0.0054	0.0112	0.0143
61	0.0154	0.0112	0.0504	0.0636
62	0.0476	0.0347	0.0523	0.0660
63	0.0731	0.0535	0.1405	0.1748
64	0.0777	0.0574	0.1031	0.1288
65	0.2339	0.1813	0.1408	0.1745
66	0.2303	0.1846	0.1843	0.2266
67	0.1416	0.1150	0.1084	0.1344
68	0.1127	0.0927	0.1413	0.1742
69	0.0961	0.0798	0.0778	0.0967
70	0.1192	0.1004	0.1522	0.1869
71	0.0617	0.0522	0.0432	0.0538

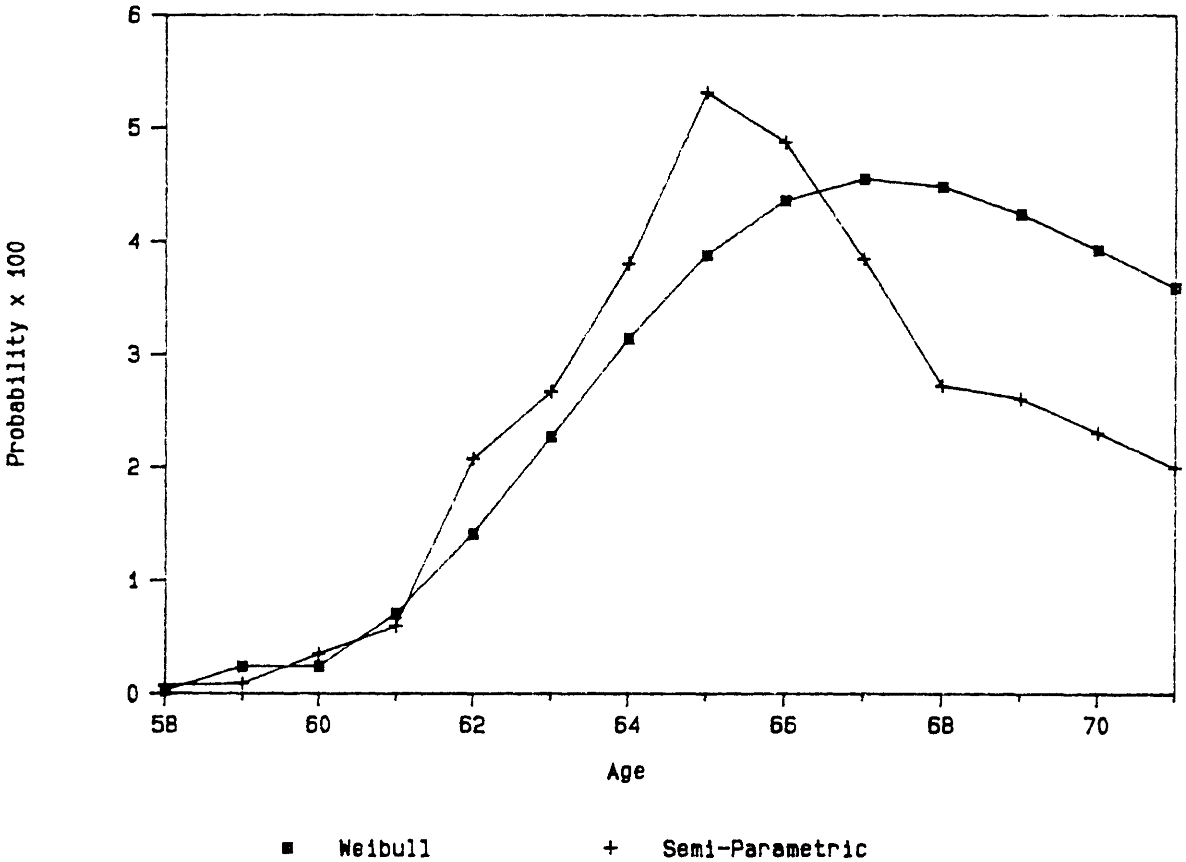
Table 9 - (continued)

Correlated Competing Risks Model

Age	Full Retirement Current Law	Full Retirement 1969 Law	Partial Retirement Current Law	Partial Retirement 1969 Law
58	0.0059	0.0036	0.0020	0.0022
59	0.0073	0.0048	0.0000	0.0000
60	0.0151	0.0106	0.0185	0.0202
61	0.0327	0.0242	0.0206	0.0222
62	0.1044	0.0830	0.0955	0.1013
63	0.1552	0.1307	0.1055	0.1108
64	0.1770	0.1544	0.2555	0.2652
65	0.4548	0.4188	0.2240	0.2318
66	0.4272	0.4016	0.3577	0.3683
67	0.3034	0.2867	0.3812	0.3922
68	0.2522	0.2395	0.2120	0.2191
69	0.2222	0.2118	0.2554	0.2642
70	0.2604	0.2495	0.1442	0.1497
71	0.1527	0.1464	0.0615	0.0615

Note: Rates are calculated by averaging the hazard rates for individuals across the 381 individuals in the youngest cohort of the sample. The calculation for 1969 law involves simulating benefits for individuals over time, given no change from the 1969 rules, and then computing the implied hazard rates.

Figure 5 - Change in Hazard Rates Resulting
From Change in Social Security Law,
Single Form of Retirement Models



5.2 Full and Partial Retirement (Competing-Risks)

Having shown that the semi-parametric model offers advantages over the commonly used Weibull form of the hazard, I now consider a more complete model of retirement behavior which accounts for the presence of partial retirement. These models are the semi-parameteric extension of the HW and DH studies to the competing risk framework.

5.2.1 Independent Risks

The previous models considered the possibility of only one form of retirement. One way to model the presence of two forms of retirement is to consider a competing risks model in which the baseline hazards are assumed to be independent. The specification is equivalent to specifying a model of the form given in (3.2.6) with independent errors. As shown by Kalbfleisch and Prentice [1980] such a model has a likelihood that factors into two separate parts, one for each hazard. The estimation technique then requires a separate estimate for each risk in which failures of the alternative type are treated as censored failures.

An independent, semi-parametric baseline hazard model is estimated for both full and partial retirement. The results for the specification with all the covariates present are given in the first two columns of table 10. These parameters can be compared directly

Table 10 - Parameter estimates for full and partial retirement models, with correlated and independent semi-parametric baseline hazards.

Variable	Independent		Correlated (2)		Correlated (3)	
	Full	Partial	Full	Partial	Full	Partial
Control Variables						
education	-0.020 (0.012)	-0.028 (0.016)	-0.034 (0.009)	-0.032 (0.010)	-0.019 (0.010)	-0.021 (0.011)
health	0.474 (0.073)	-0.012 (0.100)	0.368 (0.064)	0.268 (0.087)	0.384 (0.065)	0.220 (0.098)
married	-0.480 (0.134)	0.001 (0.178)	-0.442 (0.100)	-0.337 (0.113)	-0.443 (0.100)	-0.300 (0.117)
non-white	-0.012 (0.123)	-0.350 (0.164)	-0.089 (0.098)	-0.169 (0.099)	-0.087 (0.102)	-0.177 (0.099)
household size	0.017 (0.026)	0.010 (0.038)	0.044 (0.021)	0.037 (0.023)	0.041 (0.021)	0.034 (0.024)
mandatory retire	0.466 (0.071)	0.081 (0.101)	0.305 (0.064)	0.218 (0.082)	0.293 (0.065)	0.164 (0.086)
clerical	0.073 (0.188)	-0.029 (0.233)	-----	-----	-0.022 (0.118)	-0.064 (0.135)
craft	0.164 (0.189)	0.465 (0.228)	-----	-----	0.081 (0.124)	0.176 (0.138)
labor	0.407 (0.182)	0.115 (0.225)	-----	-----	0.221 (0.116)	0.121 (0.141)
manager	-0.097 (0.205)	-0.090 (0.266)	-----	-----	-0.218 (0.138)	-0.196 (0.155)
profess-ional	-0.208 (0.234)	0.236 (0.277)	-----	-----	-0.209 (0.159)	-0.038 (0.174)

Table 10 (continued)

Variable	Independent Full	Independent Partial	Correlated (2) Full	Correlated (2) Partial	Correlated (3) Full	Correlated (3) Partial
Economic Variables						
earnings	0.001	-0.028	-0.009	-0.014	-0.003	-0.012
	(0.010)	(0.014)	(0.007)	(0.007)	(0.008)	(0.007)
wealth	0.009	0.036	0.011	0.016	0.015	0.022
	(0.015)	(0.016)	(0.012)	(0.011)	(0.013)	(0.012)
full pension	0.111	-0.065	0.077	0.037	0.090	0.032
	(0.097)	(0.131)	(0.078)	(0.082)	(0.080)	(0.083)
part pension	-0.057	0.201	0.022	0.071	0.006	0.079
	(0.096)	(0.117)	(0.073)	(0.076)	(0.075)	(0.080)
ss62	0.587	0.084	0.411	0.310	0.411	0.257
	(0.077)	(0.100)	(0.065)	(0.078)	(0.067)	(0.085)
delta (ss65-ss62)	0.020	-0.100	-0.039	-0.061	-0.035	-0.071
	(0.015)	(0.021)	(0.012)	(0.016)	(0.013)	(0.017)
Other						
ρ	-----	-----	0.935		0.842	
			(0.119)		(0.197)	
log likelihood	-2616.14	-1799.50	-4341.15		-4319.47	

Note: Asymptotic standard errors in parentheses.

to the parameters for the single form of retirement model presented in table 7. There is considerable variation across risks in the parameters, with the single risk estimates appearing to be a combination of the competing risk estimates. Almost uniformly, the parameters for the single form of retirement model lie in the interval bounded by the full and partial retirement estimates. The parameter estimates indicate that treating both full and partial retirement as a single form of retirement convolutes the separate influences of exogenous variables upon the retirement decision.

First, notice that the results for Social Security benefits differ across the type of retirement. Additional Social Security benefits provide significant, positive incentives for full retirement but not for partial retirement. The Social Security delta reduces the conditional probabilities of partial retirement, but has little effect on the full retirement hazard. The increases in the probability of retirement resulting from larger benefits are expected, but the finding that there is variation across types of retirement is new and informative. One implication of the differential impact is that policies designed to increase benefit levels and increase marginal incentives through higher deltas will have the effect of moving some individuals from partial retirement into full retirement. The strong effect of the Social Security delta in reducing the partial retirement hazard is expected, but the lack of effect for full retirement is somewhat puzzling. This result suggests that individuals who can better adjust their behavior marginally, in the sense of being inclined to partially retire, will be more responsive to the incentive effects of Social Security. The general conclusion to be drawn from these results is that changes in the Social Security system affect the probabilities of full retirement primarily through benefit levels while affecting partial retirement through changes in the delta variable.

Furthermore, there are significant differences in the parameters for the economic variables in the competing risks case. While the earnings effects are small and imprecisely estimated in the single risk model, I find a small, but significant negative effect of earnings upon partial retirement and little effect upon full retirement. The result is a little surprising, since it is expected that earnings will have a negative effect for both forms of retirement. One possible interpretation for the stronger effect upon partial retirement is that individuals with high current earnings are more likely to be subject to the earnings test in a partial retirement job so that the high implicit marginal tax rates discourage that form of retirement. These implicit marginal tax rates are, of course, not effective for an individual who is fully retired and out of the labor force. Wealth, which is estimated imprecisely in the single-retirement model, has a small and statistically significant, positive impact upon partial retirement probabilities. The impact of additional assets upon full retirement probabilities is still unclear, as is the reason for the difference between the results for full and partial retirement.

Focusing upon the parameter estimates for the control variables, note that health status significantly increases the probability of full retirement, but not the probability of partial retirement. Similarly, being non-white reduces substantially the partial retirement hazard, but not the full retirement hazard. Both the health and non-white results make sense if one thinks of an individual receiving partial retirement job offers from some distribution. Having poor health will increase the desire for leisure, leading to higher probabilities of full and partial retirement. With poor health, however, an individual is likely to have a reduced set of partial retirement opportunities so that the probability of partially retiring should increase by a smaller amount. The different effects are not captured in structural models which assume that an individual has a single partial retirement wage offer independent of health status. In the same vein, a non-white individual is expected to have a reduced probability of partial retirement because of poorer job opportunities.

The variation across risks in the parameters for the occupation variables are consistent with the search interpretation of the partial retirement decision. Recall that in the single form of retirement model, the probability of retiring is increased for laborers. The single risk measure of the occupation effect masks variation between the effects on the full and partial retirement hazards. For the competing risks model, the conditional probability of retirement is increased for both full and partial retirement, but by a large, statistically significant amount for full retirement, and a smaller imprecisely measured amount for partial retirement. The larger effect for full retirement accords with the notion that laborers are likely to have relatively few acceptable reduced work opportunities so that full retirement is encouraged relative to partial retirement. The opposite seems to be true for craft individuals, with the partial retirement effect dominating the change in the full retirement hazard.

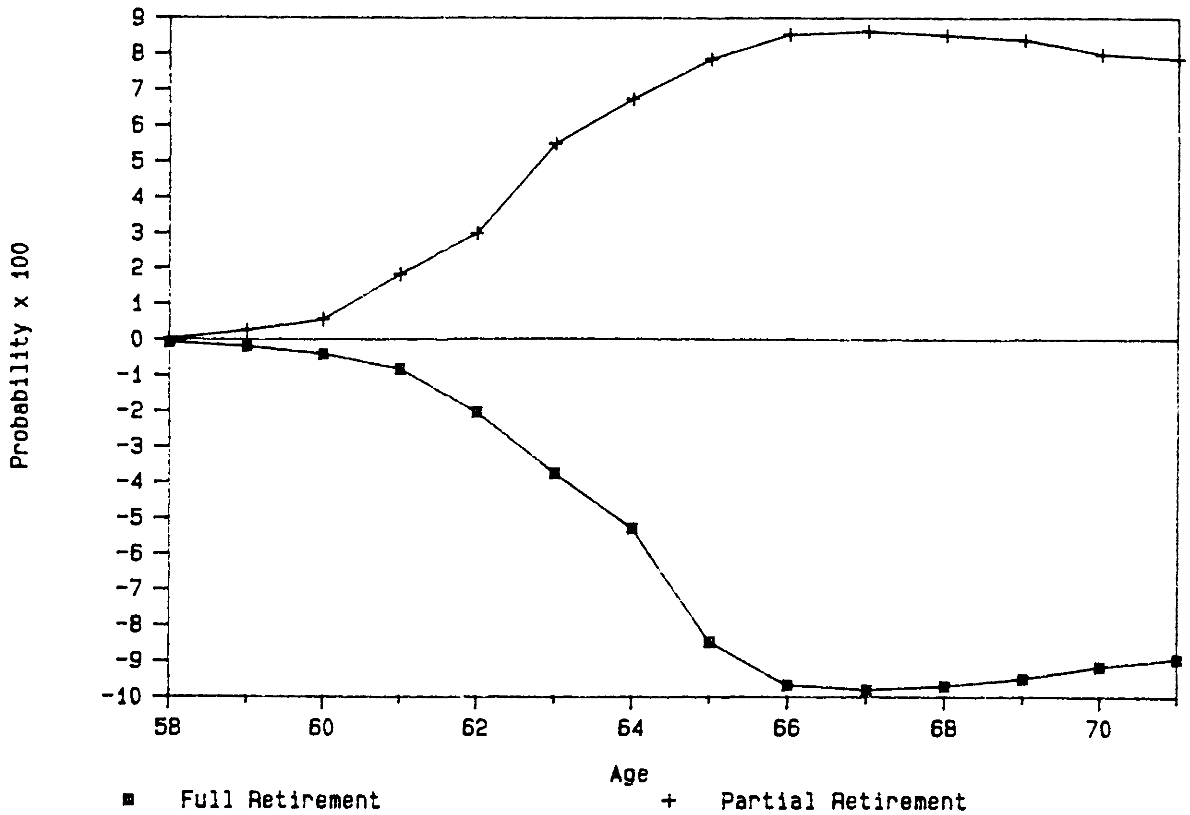
In general, the greater sensitivity of partial retirement to changes in the delta variable, as well as the different impacts of health, mandatory retirement and occupation status upon the relative retirement probabilities, suggest that a large part of full retirement behavior may not result from responses to economic incentives. Instead, the decision to retire fully may originate in physical and institutional contraints upon continued work. A case can be made for this interpretation given the strong effects of health and occupation upon the full retirement hazards.

I perform simulations to show the quantitative impact of differences across risks in the effects of Social Security variables. I again consider the survivor functions and hazard rates for the sample of individuals aged 58 in 1969 under both current and 1969 law. The simulation results for the survivor and hazard functions for full and partial retirement are presented in table 8. Figure 6 shows the changes in the "marginal" survivor functions for the two forms of retirement. The lower graph shows that the survivor function for full retirement is substantially lower under current law. For example, the sample full retirement survivor function is 5 percent lower at age 64. The difference between the survivor function under the two regimes peaks at nearly 10 percent at age 67. The upper portion of the graph shows the corresponding calculation for partial retirement. Unlike the case for full retirement, individuals have a larger partial retirement survivor function under current law. The difference ranges from 1 percent at age 59 to 9 percent at age 67. While it is difficult to generalize based upon figure 6 without taking into account the potentially latent nature of the failure times for the two forms of retirement, the results from this specification indicate that the change in Social Security law has increased the probability of early retirement (before age 65) by around 10 percent.[31]

The different simulation results for the two forms of retirement are also seen in the changes in hazard rates. The rates are

[31]To be precise, the interpretation of these changes in survivor functions as changes in labor force participation rates is not correct, since the latent nature of the risk process is ignored. These numbers do, however, provide a useful sense of the responsiveness of retirement probabilities to various factors.

Figure 6 - Change in Survivor Functions Resulting
From Change in Social Security Law,
Independent Competing-Risk Model



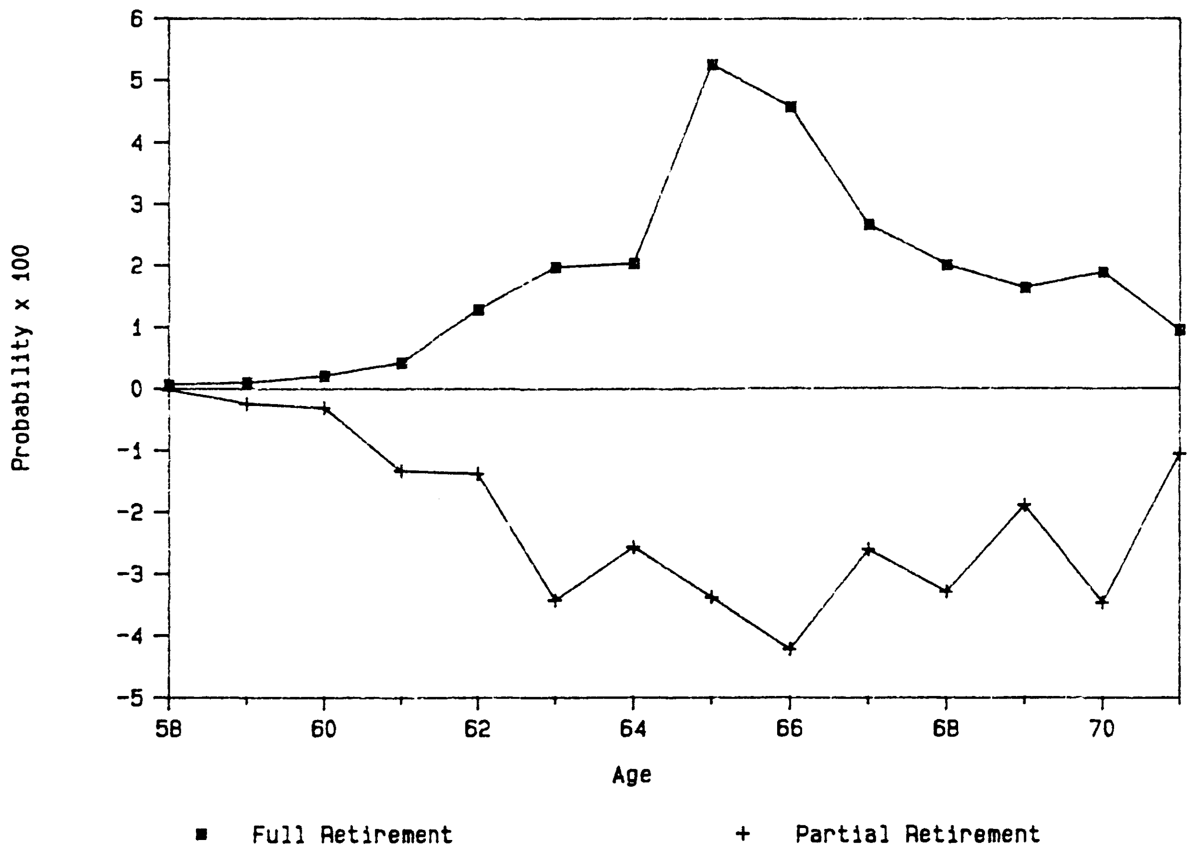
presented in table 9 and the change in rates from 1969 law to current law are depicted in figure 7. The results are reflective of the differences in parameter estimates across risks as well as the different shapes of the estimated baseline hazard rates. For the ages 62-66, the changes in the Social Security system from 1969 rules reduce the conditional probability of partial retirement by between 2 and 4 percentage points. In contrast, over the same interval, the conditional probability of full retirement increases by a slightly larger amount. At age 65 the hazard rate for full retirement is a full 5 percent higher under current law.

These results are quite striking in that they suggest that the recent changes in Social Security law have indeed altered retirement probabilities, but in different ways across retirement types. The simulation results indicate that changes in full retirement probabilities have resulted primarily through benefit levels while the partial retirement probabilities have been motivated by increases in the marginal return to continued work. While the magnitude of these effects changes under the specification of correlated risks considered in the following section, this basic result is invariant to the independence assumption.

5.2.2 Correlated Risks

The final model to be considered allows for correlation between the error terms in (3.2.6). To allow for non-independent baseline

Figure 7 - Change in Hazard Rates Resulting
From Change in Social Security Law,
Independent Competing-Risk Model



hazards, I assume that (ϵ_1,ϵ_2) are distributed as a bivariate normal with correlation coefficient ρ. Parameter estimates for the bivariate model are given in table 10. Once scaled for the unequal variances, the parameters from the bivariate normal model are comparable to the extreme value parameters for the independent competing risks model of the previous section.[32]

The estimates for ρ in table 10 are sensitive to the model specification. For the two reported specifications, ρ is 0.935 and 0.842. While estimated with considerable precision, the parameter estimate changes considerably when occupation class is considered. Nevertheless, given the magnitude of the two point estimates, it seems likely that there is a strong positive correlation between ϵ_1 and ϵ_2.[33]

An examination of the parameter estimates for the covariates reveals the importance of taking the correlation into account. Focusing on specification (3) in table 10, the estimates for the two risks are closer than the corresponding estimates for the independent

---

[32] The variance of the logit model is $\pi^2/3$. The extreme value parameters in the first two columns of table 10 should then be multiplied by 0.551 to allow for comparison with the probit estimates in the latter columns. Amemiya [1981] suggests that the logit parameters be scaled down by a factor of 0.625 since this seems to provide a better correspondence between the two models.

[33] ρ is estimated as a non-linear function of an estimated parameter. I suspect that the precision of the estimate of ρ is overstated, resulting from the impreciseness of the first-order approximation used to derive the asymptotic standard error.

competing risks model. For example, after scaling, the parameters for Social Security benefits in the independent risk case are 0.367 and 0.053 for the full retirement and partial retirement risks, respectively.[34] The corresponding parameters in the bivariate model are 0.411 and 0.257. The difference of 0.314 for the independent model is considerably larger than the latter difference of 0.154. The use of an independent risk model will therefore overstate differences between the two forms of retirement.

Furthermore, the bivariate model allows for more precise estimates of the effects of Social Security upon the alternative forms of retirement. The parameters for Social Security benefits and the Social Security delta all have the expected sign and are statistically significant. Despite the fact that the size of the differences across risks in parameter estimates are less striking, the basic character of the results of the previous section still holds: additional benefits have a stronger positive impact upon the probability of fully retiring than upon the probability of partially retiring; a larger delta reduces the partial retirement hazard by more than the full retirement hazard.

The magnitude of the effects of Social Security benefits and the Social Security delta on the relative probabilities of retirement are best seen by considering simulation results. I again use the

---

[34] For these comparisons, I use the Amemiya scaling factor described earlier.

benefits and deltas computed under 1969 law for comparison. The simulated sample survivor functions are given in table 8, the survivor function comparisons in figure 8, and the hazard rate comparisons in figure 9. The lower portion of figure 8 shows that the change in Social Security rules has reduced the full retirement survivor function. The probability that the latent failure time for full retirement is greater than 63 falls by over 5 percentage points; at age 64, the reduction is about 6 percent. These reductions are considerably smaller than the corresponding responses for the independent dual risk model. In contrast to the 10 percent increase in early retirement generated by the independent risk model, the reduction here is a little under 6 percentage points. Furthermore, the distribution of these changes differs greatly, with the change in full retirement survivor function for the bivariate model possessing a well-defined mode at age 64. The upper portion of figure 8 shows the corresponding increase in the partial retirement survivor function. Again, the responses, on the order of 1 to 2 percentage points, are considerably smaller than those for the independent model.

The simulated hazard functions in figure 9 show the same general pattern. The change in Social Security law has reduced the conditional probabilities of partial retirement, while increasing the full retirement hazards. As reflected in the survivor function simulation above, the increase in full retirement hazards ranges from

Figure 8 - Change in Survivor Functions Resulting From Change in Social Security Law, Correlated Competing-Risk Model

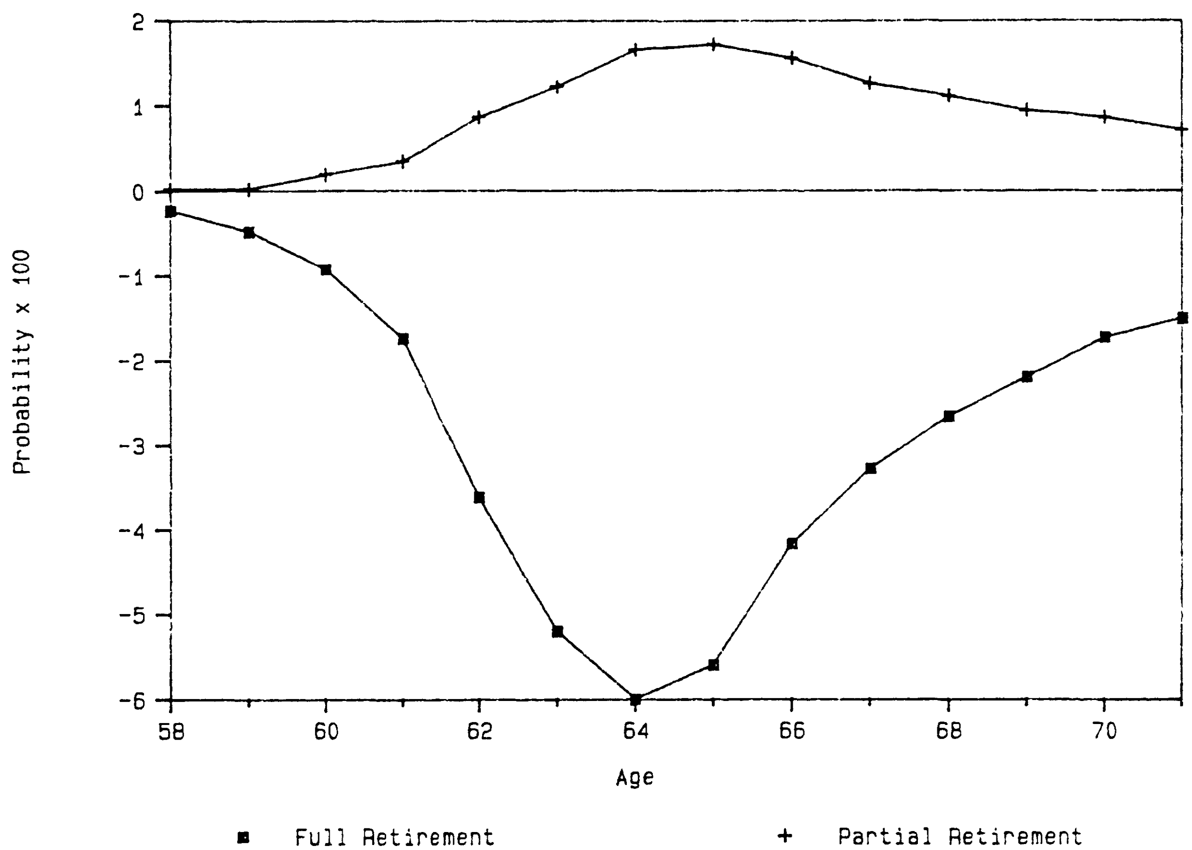
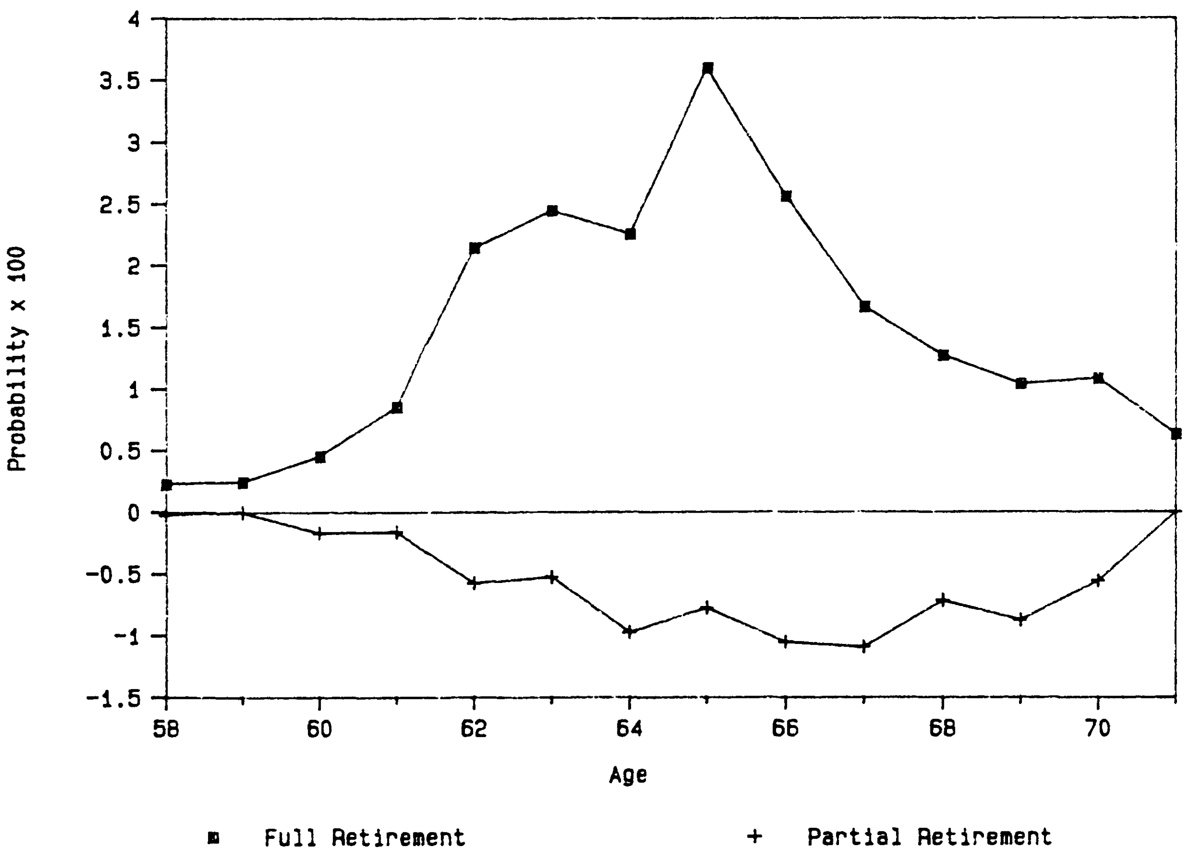


Figure 9 - Change in Hazard Rates Resulting From Change in Social Security Law, Correlated Competing-Risk Model



2 to 3.5 percentage points from age 62 to 65, a smaller change than for the independent dual risk model. Similarly, the reductions in partial retirement hazards are relatively small, about 1 percent versus the 2-4 percent changes derived under the independence assumption.[35]

The larger predicted effects under the assumption of independent risks seem to result directly from the lack of accounting for the correlation between risks. As an illustration, note that the survivor function estimates in table 8 for the two models are very different, especially in the upper tail of the age-distribution. The differences in the survivors can be attributed to the fact that the independent model treats failure of alternative types as censored observations when calculating the risk for a given failure. In contrast, the correlated model uses the information about the correlation to more precisely calculate the true probabilities of survival, in essence using the information that partial retirement is more like full retirement than it is like the censoring category.

Finally, the simulation results for the various models are summarized in table 11 where the percentage change in early retirement is presented for both the single and dual risk models. Notice first that the predicted responses of full retirement to the

---

[35] I also calculated the minimum χ^2 test for the two semi-parametric baseline hazards against the two-parameter Weibull specifications. The resulting value of 271 is larger than the 5% critical value for $k_1+k_2-4 = 24$ *df*.

Table 11 - Estimated changes in survivor probabilities and hazard rates resulting from changes from 1969 Social Security law, various models and ages, 381 individuals.

	Age 62	Age 65	Age 70
	Change in Survivor Probability x 100		
<u>Single Risk</u>			
Weibull			
Semi-Parametric	-4.056	-7.636	-1.456
(Extreme)	-2.921	-6.708	-2.163
<u>Dual Risk</u>			
Independent			
Full	-2.035	-8.476	-9.153
Partial	2.973	7.842	7.984
Correlated			
Full	-3.603	-5.596	-1.726
Partial	0.863	1.722	0.861
	Change in Hazard Rates x 100		
<u>Single Risk</u>			
Weibull	1.416	3.875	3.926
Semi-Parametric	2.077	5.312	2.307
(Extreme)			
<u>Dual Risk</u>			
Independent			
Full	1.284	5.259	1.883
Partial	-1.372	-3.379	-3.466
Correlated			
Full	2.145	3.602	1.085
Partial	-0.577	-0.777	-0.557

Note: These changes are calculated from the numbers presented in tables 8 and 9. The survivor probabilities are computed by applying the estimated baseline hazard for the model is question and the estimated β to actual individual-level data and to data simulated on the basis of 1969 law. The resulting survivor functions and hazard rates are then averaged over the youngest RHS cohort.

Social Security reforms fall across specifications, first as the Weibull specification is relaxed, and then as partial retirement, and correlated partial retirement are considered. A similar result is observed for the response of partial retirement when the independence assumption is removed.

Conclusion

I will not dwell excessively upon the results since they have been already been discussed at some length. In short, I find that Social Security affects full and partial retirement in different ways. As a result, the recent increases in Social Security benefits are shown to have increased the probability of full retirement and reduced the probability of partial retirement. The semi-parametric competing risks model employed in this study appears to be superior to previously used duration models. The relative absence of functional form restrictions yields predicted changes in retirement probabilities that are smaller than those estimated with restrictions.

The finding that partial retirement behavior differs substantively from full retirement behavior is of particular interest since, given current trends in mortality and morbidity, partial retirement is likely to become of greater importance in the future. As the general health level of the elderly rises and institutional constraints upon work by the elderly are relaxed, delayed and partial

Chapter 2

Unemployment Insurance Benefits and the Duration of Unemployment Spells: Evidence from the Survey of Income and Program Participation

Introduction

The factors that influence unemployment have, for some time, been the subject of considerable academic scrutiny. This attention has touched on a variety of topics but has, in recent years, focused upon two issues: the relevance of a search interpretation in explaining unemployment durations, and the role of unemployment insurance (UI) benefits in altering the length and frequency of unemployment spells. Despite advances in both economic theory and econometric technique, much of the existing empirical work has been handicapped by data limitations which hamper the simultaneous study of these issues.

In this paper, I address jointly the issues of the impact of UI benefits upon exit from unemployment and the forms that exit takes. The primary contribution of this chapter is the development of a new sample of unemployment spells from the Survey of Income and Program Participation (SIPP) which allows for the estimation of a competing risk model of exit from unemployment in which individuals either are recalled to their previous employer or find a new job. The use of the SIPP in this study solves a number of the data problems that have plagued previous researchers and provides a platform for further work in the analysis of the nature of unemployment.

I find that unemployment benefits generally have a significant and precisely measured impact upon exit rates from unemployment. The basic result that UI is negatively related to hazards holds whether unemployment spells are analyzed in a single or in a competing risk framework and whether UI benefits are represented by an indicator variable, by replacement rates, or by actual benefit levels. In the competing risks framework, the effects differ across risks with the recall hazard more sensitive to UI. I also find that the use of measures which incorporate information about the extent of benefits improve the explanatory power of the model relative to the use of receipt indicators.

Furthermore, the results suggest that UI benefit exhaustion has important effects upon the exit hazard. In particular, I find evidence of a spike in the exit hazard at 39 weeks. Given the precision of the SIPP weekly data on employment status and the presence of spikes in the estimated hazard at exhaustion points, I conclude that the increases in exit hazards coincident with benefit exhaustion points are not the result of discrete responses by individuals to survey questions.

There is mixed support in the data for a search model of new job findings. While the overall hazard rate is estimated to be a decreasing function of duration, the new job hazard for exit from unemployment appears to be increasing over time, at least for the first few months. The shape of this hazard is estimated with considerable error so that strong conclusions should not be drawn based upon these results. Nevertheless, the finding of an increasing new job hazard supports the findings of Katz [1986] and Han and Hausman [1986] for a sample from a different survey. On the other hand, the large errors associated with the estimates, the non-monotonicity of the hazard, and the relatively small number of new job findings in the data suggest that the findings of increasing hazards from these previous studies may not be robust across alternative data sources.

More generally, the results from the SIPP suggest that the recent concern with the duration of a single unemployment spell may be somewhat missplaced. A large percentage of the individuals who experience a spell of unemployment during the 16 month period experience multiple spells. This result suggests that a more general analysis of durations that incorporates information about spell frequency is needed.

The remainder of the chapter is divided into four sections. In section 1, I outline some of the basic issues in the study of unemployment spells. The discussion is brief since much of this material should be quite familiar. In sections 2 and 3, the development of the SIPP data set is discussed and descriptive statistics are presented. The latter section also includes the empirical hazard rates for exit from unemployment. This discussion is followed by results from both single risk, semi-parametric estimation and competing risk models. There is a brief conclusion in section 4.

1. Background

The analysis of unemployment durations has proceeded along two related dimensions. First, many analysts have focused their efforts on examining the nature of duration dependence; more specifically, on determining whether hazards are increasing or decreasing with duration length. Also of interest has been the issue of whether state unemployment insurance (UI) systems provide incentives for individuals to prolong job search, thereby increasing the pool of unemployed individuals and observable unemployment.

Much of the recent empirical literature has focused upon reconciling observed hazard rates with the theoretical framework of job search models. At the center of the discrepancy is the fact that reemployment hazards are invariably estimated to be decreasing over time. The decreasing hazards are in direct contrast to the predictions of standard models of job search (McCall [1970], Mortensen [1977], Burdett and Mortensen [1980], Pissarides [1982]). In these models, the optimal individual search strategy is to set a reservation wage and to compare arriving job offers to this reservation wage. It can be shown that the reservation wage will, for a variety of reasons, decrease over time. If the wage-offer distribution is stationary, then the result will be an increasing hazard rate.

It has been suggested that the observed decreasing hazard rate is explained by the presence of unobserved individual heterogeneity. Heckman and Singer [1985] demonstrate that heterogeneity will unambiguously bias the hazard rate downward, with larger biases occurring over time. The intuition behind the direction of the bias is that individuals with relatively good draws from the heterogeneity distribution are more likely to find employment early so that as time passes, the sample selection process concentrates the sample toward those less likely to find jobs.

An alternative explanation has been suggested by Katz [1986]. He argues that the problem results from an ambiguous definition of reemployment. Katz notes that for individuals on layoff, there are two potential forms of exit from unemployment: one in which individuals find a new job, and another in which they are recalled to their previous employer. He shows, using retrospective data on unemployment spells, that if one estimates a competing risks, proportional hazards model of new job and recall risk, the overall hazard can be split up into two parts, with the decreasing portion of the overall hazard coming from the layoff component, and with the new job hazard increasing over time as predicted by theory.

Despite these and other important advances in the study of hazard models of unemployment, it is important to note that the duration framework is ill-suited to an examination of the causes of unemployment. The basic hazard framework focuses attention on the narrow question of what factors influence the length of a spell, ignoring possible interactions between spell length and frequency. This emphasis on chronic unemployment has persisted despite evidence that much unemployment is related to employment instability (Hall [1972], Marston [1975], Clark and Summers [1976]).

A second branch of the unemployment literature has examined the extent to which state UI insurance schemes have increased the frequency and duration of unemployment spells. Feldstein [1978] for one, has argued that the existence of UI provides a subsidy to search, thereby increasing the duration of any given unemployment spell. Furthermore, Topel [1983] has noted that the increase in observed unemployment resulting from UI has two components, one coming from increased durations, and the other from more frequent layoffs resulting from imperfect firm experience ratings. The imperfect ratings imply that the marginal benefit accruing to firms from the layoff of an additional worker is not equal to the increased insurance cost that is incurred. Firms with relatively high unemployment levels may find themselves in a situation where benefit payments exceed the statutory maximum on tax liabilities so that the additional payment of benefits is associated with zero additional liability.

The relationship between individual search behavior and the payment of UI should have empirically verifiable effects. First, UI benefits should be positively related to spell durations. All things equal, individuals who receive benefits are facing a lower price, in the form of forgone wages, to continuing search. The subsidy implies that UI benefits will increase the duration of a given spell by increasing the reservation wage. It also implies that the hazard rate for exit from unemployment should increase upon UI benefit exhaustion. Additionally, since the receipt of UI benefits is directly related to an individual's labor market state, UI represents a direct subsidy to search so that a dollar of UI benefits should have a stronger effect in prolonging durations than a corresponding dollar from other income sources.

The positive impact of UI upon the average duration of an unemployment spell is mitigated to a degree if firms with imperfect experience ratings use the UI system to subsidize layoff. Since firms are not paying the full insurance cost of generating unemployment, they will tend to layoff more frequently than they would otherwise. If these spells of unemployment are shorter than the existing spells then the presence of a UI system might appear to be related to shorter spells, even though aggregate unemployment has risen. While the current study does not address the question of what determines the frequency of unemployment spells, the impact of imperfect experience rating differentials across industries should yield an observable industry component to spell duration. More specifically, one should observe some industries with high unemployment rates that also have high unemployment exit hazards.

2. *Data Issues*

The data on unemployment durations used in this paper are derived from waves 1 through 4 of the Survey of Income and Program Participation (SIPP). Administered by the Bureau of the Census, the SIPP is a longitudinal survey of approximately 21,000 household units selected to be representative of the non-institutional population of the United States. The survey is designed to provide information on income, labor force activity and participation in a variety of transfer programs at the individual, family, and household level.

Over the course of a two and one-half year period, individuals are sampled once every four months for a total of 9 waves of data. Waves 1 through 4 of the SIPP cover the period from October 1983 to December 1984, consequently, data for a given individual in this study are available for at most 16 months. In each survey period, questions on income receipt and labor force status are asked pertaining to the four months preceeding the interview date.

Since the accuracy of the spells used in this analysis is directly related to the reliability of the data both within a given SIPP wave and across the panel, it is important to note that the data appear to be of high quality. The designers of the survey have taken particular care to insure that attrition is kept to a minimum and that the data are relatively free from error. If necessary, individuals who move are followed to new addresses. Most questioning is conducted via personal visits subject to geographic restrictions.[1] Callbacks and consistency checks are conducted to minimize the amount of missing or erroneous data.

The analysis in this paper is carried out at the level of heads of household with individuals matched across successive waves of the survey. Because of differences in the sampling procedure used for wave 2, only about 3/4 of the original 21,000 households were candidates for matching.[2] There were 12,707 matches across all four waves. Of those matches, 5,546 spells of unemployment were observed, of which 4,222 were completed duration spells. Complete and incomplete spells are used in this study. Since one of the focal

---

[1]Individuals who move and are not located within 100 miles of one of the original sampling areas are subsequently interviewed by telephone. Though I have not tabulated the sample by type of interview, given the scope of the original survey, the number of telephone interviews is likely to be small. Many other samples, including, in particular, the Current Population Survey, do not follow movers.

[2]The SIPP is divided into four rotation groups which are interviewed on successive months. A typical SIPP "wave" is made up of interviews for all four rotation groups. Wave 2, however, contains only records for individuals in the first 3 rotation groups. This unbalanced sampling procedure was designed so that specific questions on personal taxes to be asked in wave 6 could be asked in the months of May, June, July and August.

points of this analysis is on determining the mode of exit from unemployment, individuals for whom employer information was not available upon completion of spell were excluded from the final sample. Also excluded were individuals for whom job information at the onset of the spell was not available. The resulting data set consists of 4,224 spells, censored and uncensored, distributed among 2,282 individuals.

There are a number of reasons that the SIPP is a particularly valuable source of information about an individual's labor force status. Previous analyses of labor force transitions have suffered from restrictions on the availability of key data. While not a perfect data source, the SIPP addresses the most serious of the difficulties.[3] The following discussion focuses on the problems associated with a variety of data sources and the ways in which the SIPP allows one to address the most important of the difficulties.

Despite a variety of well-documented problems, the Current Population Survey (CPS) is the most widely used source of information on labor force transitions. Prior to the introduction of the SIPP, the CPS was unique in its universal coverage of the working population. For the study of unemployment durations, however, the

---

[3]Moreover, the expected availability of the wave 3 supplement which contains work history information promises to make the current SIPP panel a platform for the analysis of issues that cannot be addressed given current data. In this regard, it is the combination of the wealth of data and the universal coverage that is of particular value.

monthly CPS data are severely hampered by the use of a point-in-time sampling technique whereby spells of unemployment are only observed while in progress. As noted by a number of researchers, this sampling procedure results in a sample which understates the frequency of shorter spells since longer spells are more likely to be observed.[4] More importantly for the analysis of durations, Poterba and Summers [1986] have suggested that there may be serious reporting error across successive months of the CPS leading to an overstatement of the frequencies of labor market transitions.

Other sources of data such as the Panel Study of Income Dynamics (PSID) used by Katz [1986] and Han and Hausman [1986] avoid the problems associated with the point-in-time technique by obtaining retrospective information on unemployment experience. Individuals are asked to report on the dates associated with the start of their recent unemployment spells and the duration of the spells. Also recorded is the individual's current labor market status. From this, the duration and status (censored or uncensored) of unemployment spells can be computed. The PSID also allows for the identification of whether a spell of unemployment ends in a new job or in recall.

The use of retrospective surveys introduces the possibility of error in reported durations resulting from discreteness of responses

---

[4]This discussion ignores the CPS rotation group bias. The problems with the point-in-time technique are discussed in more detai in Kaitz [1970], Marston [1975] and Salant [1977]. See also Kiefer, Lundberg and Neumann [1984].

to survey questions. It is easy to imagine individuals reporting that their unemployment spells lasted 1/2 or 3/4 of the previous year. This discreteness has non-trivial consequences since one important finding from studies using retrospective surveys is that the hazard rate of exit from unemployment increases sharply at 26 and 39 weeks. A number of researchers have noted that this increase in the hazard rate coincides with the typical exhaustion points of state unemployment insurance benefits. Unfortunately, whether benefit exhaustion is causal or whether individuals are simply reporting durations in fractions of a year is difficult to determine given retrospective data.

A different type of data drawn from state unemployment agency administrative records is employed by Moffitt [1985] and Meyer [1987]. These data have the advantage of providing continuous sampling of individual unemployment spells on a weekly basis and therefore provide a more accurate depiction of the length of spells than either the CPS or the PSID data. Unlike the PSID, the administrative record data contain information on the amount of unemployment insurance benefits that an individual receives. By their very nature, however, administrative records are available only for those individuals who receive unemployment insurance benefits. Given the high fraction of the unemployed who do not receive benefits, such a sample is clearly not representative of the population of unemployed individuals. Furthermore, administrative records typically cannot be used to determine whether exit from unemployment is via recall or a new job. Katz's study of PSID spells suggests that this is an essential piece of information for attempting to understand the nature of unemployment durations.

The SIPP provides information which addresses many of the problems associated with existing duration data. For one, the SIPP provides a unique opportunity to examine a large, representative sample of the working population. This includes both individuals who receive and those who do not receive unemployment insurance benefits. The data also support an examination of the influence of UI benefit levels and the importance of layoff versus recall as a form of exit. While the UI benefit data are survey values and therefore less precise than values obtained from administrative records, the presence of employer information in the SIPP allows one to examine benefits in the context of a competing risks framework. This type of analysis cannot be undertaken with standard administrative records. Conversely, while the PSID allows for the identification of individuals who find new jobs, it cannot readily be used to examine the impact of UI benefit levels.

The survey design also facilitates the study of unemployment durations by providing weekly data on employment status. The SIPP employs a continuous monitoring technique in which individuals are asked to report on their labor market status in each week of the preceeding four months. Since individuals are asked to report on their status for each week, the point-in-time spell observation problems associated with the CPS do not exist. The combination of a relatively short four-month interview period and continuous monitoring should make the data less prone to rounding error of the type expected in yearly surveys like the PSID.

Central to the current analysis is the calculation of spells of unemployment. Spells for individuals in the SIPP are calculated in a straightforward fashion. When a week of paid work is followed by either a week without a job or a week with a job but without pay, an unemployment spell is said to have begun. At the time the spell begins, relevant economic and demographic data are recorded; thus the covariates used in this paper are start of spell values. For example, the asset income data is the monthly receipt of asset or property income prior to spell onset. Also recorded at the start of the spell is the SIPP employer code.

Individuals are followed from week to week across waves for the duration of the spell or until the end of the fourth wave. A spell is completed when the individual reports being with a job and receiving pay during a given week. No attempt is made separately to allow for spell termination via exit from the labor force. Censored spells are marked accordingly. For individuals with completed spells, the employer identifier at the new job is compared with the original employer identifiers; non-matching observations are recorded as new job exits.[5]

3. *Results*

3.1 *Descriptive Results*

As noted previously, the recorded covariates in the SIPP data used in this study are start of spell values. Conceptually, the covariates analyzed in this paper can be divided into three classes: heterogeneity controls, income variables, and industry and occupation classification variables. The heterogeneity controls are standard individual characteristics such as age and level of education. Also included are indicator variables for being female, non-white and married. The industry and occupation class variables are based upon 1980 Census of Occupation and Industrial Classification groupings, and are included to control for differences in labor market structures and opportunities. Frequencies for the industrial and occupation classes are presented in table 1.

Two types of income variables are used. The first type is related to the amount of asset income that the family receives. If the search model interpretation of unemployment is correct, individuals with high asset income values would be expected to have

---

[5]This is essentially the same technique used by Katz in his PSID study. A more detailed classification should be possible with the availability of the SIPP work history supplement. Tenure information available in the supplement can be used to check the consistency of responses.

Table 1 - Industry and occupation classification frequencies for the SIPP sample of unemployment spells, 1980 Census of Industry and Occupation Classification System.

Occupation	*Number*	*Frequency*
Professional	566	13.40%
Sales	838	19.84%
Service	676	16.00%
Craft	888	21.02%
Laborer	1256	29.73%
Total	4224	100.00%

Industry	*Number*	*Frequency*
Agriculture	18	0.43%
Mining	47	1.11%
Construction	634	15.01%
Durables	408	9.66%
Non-durables	707	16.74%
Transportation	370	8.76%
Wholesale Durables	63	1.49%
Wholesale Non-durables	584	13.83%
Services	1363	32.27%
Other	30	0.71%
Total	4224	100.00%

Note: Based upon author's calculations from the Survey of Income and Program Participation.

longer durations since the marginal value of foregone earned income will be high for an individual with no other income sources. The other class of income variable relates to the receipt of UI. Three proxies for UI are used in estimation: an indicator variable for receipt of benefits, a measure of the UI replacement rate which compares UI benefit levels to previous earned income, and UI benefit levels themselves.

In table 2, I present descriptive statistics for the sample of 4,224 unemployment spells. The mean value of duration is presented without correction for right censoring. Furthermore, the UI figures reported in the table do not correct for the large number of individuals who do not receive UI benefits.[6] While the receipt of UI benefits has been a focal point of interest in unemployment durations, often overlooked is the fact that there are large numbers of individuals do not receive UI benefits; in the SIPP data, over 75 percent of the spells involve no receipt. Whether the lack of benefit receipt occurs because individuals do not qualify or because the majority of spells are so short that applying for and receiving benefits is not worth the effort is a question for which the answer

---

[6]The appropriate technique for correcting mean values for censoring is a subject of some debate (Marston [1975]). Since longer spells are more likely to be censored than short, a correction would merely accentuate the differences between mean values. For the full 4,224 spell sample, the mean values conditional upon receipt are $177.97 for benefits and 0.327 for the replacement rate. The standard deviations are 113.43 and 0.472 respectively.

Table 2 - Mean values for covariates in the SIPP sample of unemployment spells, broken down by type of exit.

Variable	*Overall* Mean	*Overall* Std Dev	*Recall* Mean	*Recall* Std Dev	*New Job* Mean	*New Job* Std Dev
Duration	10.45	14.32	5.78	8.22	10.57	9.57
Heterogeneity Controls						
Age	41.76	14.42	40.48	13.63	34.22	10.46
Education	12.11	2.91	12.16	2.87	12.88	2.57
Female	0.33	0.47	0.32	0.47	0.32	0.47
Married	0.59	0.49	0.61	0.49	0.52	0.50
Non-white	0.14	0.35	0.14	0.34	0.07	0.25
# of children	0.85	1.17	0.86	1.16	0.91	1.38
Income Variables						
Asset Income $	68.88	270.09	65.60	264.11	40.50	120.75
UI Receipt	0.22	0.41	0.23	0.42	0.25	0.43
UI Replacement	0.06	1.18	0.06	0.25	0.10	0.30
UI Benefits $	34.63	86.41	34.08	84.03	49.05	97.66
N	4224		3127		152	

Note: Based upon author's calculations from the Survey of Income and Program Participation.

is not known and one which deserves further study.[7]

The most obvious feature of the mean values in table 2 is the difference between mean duration for spells ending in recall and the mean duration for those ending in a new job. While the sample is predominantly made up of short recall spells, the combination of the relatively lengthy new job and censored durations is enough to raise the overall mean to a figure in excess of two months. This is comparable to the average for new job spells and almost twice as great as that for recall. The figures in table 2 also suggest that new job spells are associated with younger, more educated individuals. There is nothing surprising about this relationship since job mobility is typically associated with those characteristics.[8]

---

[7]An additional possibility is that there are delays between the onset of unemployment spells and the receipt of benefits so that the actual receipt of benefits is missed. Baily [1978], for one, points out that benefits are not typically paid for the first week of unemployment. Thus, for a sample with an extremely large number of short spells, a large number of spells may appear to have no benefits associated with them. It is unlikely tht this effect accounts for the bulk of non-recipients. As one piece of evidence, note that the large proportion of non-recipients holds whether actual receipt of benefits or self-reported eligibility is used. The latter meausre should be less sensitive to timing considerations.

[8]While I do not report the standard errors of the means, they are small enough so that the comparisons are valid. For example, the standard error of the mean age for spells ending in a new job is 0.849; for education, the corresponding figure is 0.208. The standard errors are even smaller for the much larger sample of spells ending in recall.

In table 3 I present the calculated durations for the sample. Bear in mind that the 4,224 exit times include multiple observations on individuals. These durations are broken down by mode of exit, whether via a new job or recall to a previous employer. The censoring category refers to the right censored observations for whom spells were continuing at the conclusion of the sample. The same information is presented in table 4 for a subsample formed by taking the observation on the longest spell for each of the 2,282 individuals. If individuals are more likely to report longer spells, this is the type of sample that might be observed upon asking individuals retrospectively to report on their recent unemployment experience.

There is evidence that even the weekly data in the SIPP generates some discreteness in calculated durations. In table 3, the number of censored observations is noticeably elevated around 52 weeks and to a lesser degree, around 39 and 26 weeks. Since these figures are for censored observations, they result presumably not from discrete responses on the part of individuals, but because of the four-months per wave sample design of the SIPP. Significantly, however, there is no elevation of either the recall or the new job figures at the 52 week point so that discretness in censoring does not appear to be mirrored in the other exit modes.

One striking feature of the exit types in table 3 is the relatively small number of new job exits. Katz observed

Table 3 - Calculated durations of unemployment spells for the SIPP sample, broken down by mode of exit. Sample of 4,224 spells.

Week	*Risk Set*	*Total*	*Non-Censored*	*Recall*	*New Job*	*Censored*
1	4224	1153	1116	1101	15	37
2	3071	584	545	536	9	39
3	2487	325	294	284	10	31
4	2162	246	216	201	15	30
5	1916	164	141	130	11	23
6	1752	118	99	92	7	19
7	1634	122	92	85	7	30
8	1512	86	68	62	6	18
9	1426	94	72	65	7	22
10	1332	72	58	53	5	14
11	1260	66	50	43	7	16
12	1194	61	46	40	6	15
13	1133	73	51	46	5	22
14	1060	50	35	34	1	15
15	1010	63	54	48	6	9
16	947	16	14	12	2	2
17	931	124	63	57	6	61
18	807	115	35	30	5	80
19	692	12	6	5	1	6
20	680	24	18	16	2	6
21	656	20	12	11	1	8
22	636	37	22	20	2	15
23	599	21	15	13	2	6
24	578	14	8	8	0	6
25	564	18	9	7	2	9
26	546	25	14	13	1	11
27	521	12	6	5	1	6
28	509	22	10	9	1	12
29	487	18	10	8	2	8
30	469	7	4	4	0	3
31	462	14	6	5	1	8
32	448	13	9	8	1	4
33	435	7	3	2	1	4
34	428	21	13	12	1	8
35	407	71	12	12	0	59
36	336	9	2	2	0	7
37	327	7	4	4	0	3
38	320	7	1	1	0	6
39	313	18	5	5	0	13
40	295	14	6	6	0	8

Table 3 - (continued)

Week	*Risk Set*	*Total*	*Non-Censored*	*Recall*	*New Job*	*Censored*
41	281	4	1	1	0	3
42	277	9	5	4	1	4
43	268	5	2	2	0	3
44	263	10	4	4	0	6
45	253	5	2	1	1	3
46	248	8	1	1	0	7
47	240	3	1	1	0	2
48	237	10	3	3	0	7
49	227	4	1	1	0	3
50	223	7	2	2	0	5
51	216	5	3	3	0	2
52	211	88	1	1	0	87
53	123	57	2	2	0	55
54	66	7	2	2	0	5
55	59	5	1	1	0	4
56	54	6	1	1	0	5
57	48	4	0	0	0	4
58	44	5	1	0	1	4
59	39	3	0	0	0	3
60	36	4	0	0	0	4
61	32	7	0	0	0	7
62	25	1	0	0	0	1
63	24	1	0	0	0	1
64	23	6	2	2	0	4
65	17	4	0	0	0	4
66	13	2	0	0	0	2
67	11	4	0	0	0	4
68	7	4	0	0	0	4
69	3	3	0	0	0	3
Totals		4224	3229	3127	152	945

Note: Based upon author's calculations from the Survey of Income and Program Participation.

approximately twice as many new job findings in a sample a quarter of this size. One possible explanation for the difference lies in the high concentration of non-whites and industrial workers and under-representation of other groups in his PSID sample. For example, approximately 17 percent of the PSID sample is comprised of individuals in construction, a much smaller figure than the corresponding 32 percent in the SIPP shown in table 1. Similarly, laborers account for 50 percent of the PSID and professionals less than 4 percent. The figures for the SIPP are 30 and 13 respectively. Given the presumption that job change rates differ by industry and occupation, then the sampling differences between the two studies might account for the discrepancy in relative magnitudes.[9]

Also of note in table 3 are the large number of short duration spells. For the full sample, almost half of spells are observed to end in the first month, with 2,171 recalls or new jobs occurring during that period. By the end of the fourth month, the risk set has been reduced to 947. While this figure does not correct for the censored failures over the four month interval, that does not change the basic result that most spells are of very short duration.

The existence of this type of employment instability relates

---

[9]It may also be that the algorithm used here to determine whether the individual returns to the old job or finds a new one misses some job changes. Alternatively, surveys employing retrospective questioning may miss spells of short duration that end by recall because individuals do not bother to report them.

Table 4 - Calculated durations of unemployment spells for the SIPP sample, broken down by mode of exit. Subsample of 2282 spells.

Week	*Risk Set*	*Total*	*Non-Censored*	*Recall*	*New Job*	*Censored*
1	2282	470	456	450	6	14
2	1812	286	274	272	2	12
3	1526	177	166	163	3	11
4	1349	132	119	115	4	13
5	1217	87	76	73	3	11
6	1130	62	56	53	3	6
7	1068	72	56	54	2	16
8	996	48	38	38	0	10
9	948	54	42	38	4	12
10	894	43	34	32	2	9
11	851	39	32	29	3	7
12	812	40	30	26	4	10
13	772	44	33	31	2	11
14	728	21	17	17	0	4
15	707	41	34	32	2	7
16	666	10	9	9	0	1
17	656	87	49	45	4	38
18	569	74	25	22	3	49
19	495	6	3	3	0	3
20	489	16	13	12	1	3
21	473	15	9	8	1	6
22	458	24	15	13	2	9
23	434	9	6	6	0	3
24	425	10	6	6	0	4
25	415	11	4	4	0	7
26	404	17	9	8	1	8
27	387	9	6	5	1	3
28	378	11	6	6	0	5
29	367	11	7	5	2	4
30	356	6	3	3	0	3
31	350	8	3	2	1	5
32	342	9	6	6	0	3
33	333	3	1	1	0	2
34	330	20	12	11	1	8
35	310	48	8	8	0	40
36	262	7	2	2	0	5
37	255	5	3	3	0	2
38	250	4	1	1	0	3
39	246	16	4	4	0	12
40	230	10	5	5	0	5

Table 4 - (continued)

Week	*Risk Set*	*Total*	*Non-Censored*	*Recall*	*New Job*	*Censored*
41	220	4	1	1	0	3
42	216	9	5	4	1	4
43	207	2	1	1	0	1
44	205	7	3	3	0	4
45	198	2	1	1	0	1
46	196	7	1	1	0	6
47	189	3	1	1	0	2
48	186	10	3	3	0	7
49	176	3	1	1	0	2
50	173	5	1	1	0	4
51	168	5	3	3	0	2
52	163	72	1	1	0	71
53	91	42	1	1	0	41
54	49	2	1	1	0	1
55	47	4	1	1	0	3
56	43	5	1	1	0	4
57	38	1	0	0	0	1
58	37	4	1	0	1	3
59	33	3	0	0	0	3
60	30	4	0	0	0	4
61	26	5	0	0	0	5
62	21	1	0	0	0	1
63	20	3	0	0	0	3
64	17	5	2	2	0	3
65	12	1	0	0	0	1
66	11	4	0	0	0	4
67	7	4	0	0	0	4
68	3	3	0	0	0	3
Totals		2282	1707	1648	59	575

Note: Based upon author's calculations from the Survey of Income and Program Participation.

back to earlier questions about the nature of unemployment; whether unemployment is the result of chronic unemployment spells or frequent short spells. Along these lines, the difference in the numbers of spells represented by the two samples in tables 3 and 4 deserves some comment. The totals imply that individuals in the SIPP who reported spells of unemployment had an average of close to two spells. The distribution of numbers of spells is presented in table 5. While the majority of individuals experienced just one spell, fully 40 percent of had multiple spells in the 16 month period, with over 20 percent experiencing three or more spells. This large proportion of the sample with multiple spells indicates that despite the relatively high exit rate from unemployment in the first few weeks of a given spell, individuals may be experiencing relatively lengthy aggregate spells of unemployment over any given period of time.

Associated with the distribution of exit times given in tables 3 and 4 are hazard rates for exit from unemployment. These hazards are calculated by taking the number of exits of a given type at a given period and dividing by the risk set of individuals who have not yet exited prior to the start of the period. Monthly hazards for the full sample are presented in figures 1 and 2 for both exit via recall and exit via new job. The recall hazard has the characteristic downward slope found in most empirical hazards. There are small spikes in the hazard at 5 months and at 9 months. The latter spike is the familiar one associated with 36 weeks and may be related to

Table 5 - Frequency distribution for the number of spells per individual in the SIPP unemployment data set.

Number	*Count*	*Percent*
1	1625	58.29%
2	602	21.59%
3	246	8.82%
4	137	4.91%
5	79	2.83%
6	45	1.61%
7	21	0.75%
8	18	0.65%
9	12	0.43%
10	1	0.04%
11	1	0.04%
12	1	0.04%
Total	2788	100.00%

Note: This table was calculated from a slightly different sample than that used in estimation. Based upon author's calculations from the Survey of Income and Program Participation.

Figure 1 — Recall Hazard for Exit from Unemployment

Full 4,224 Spell Sample

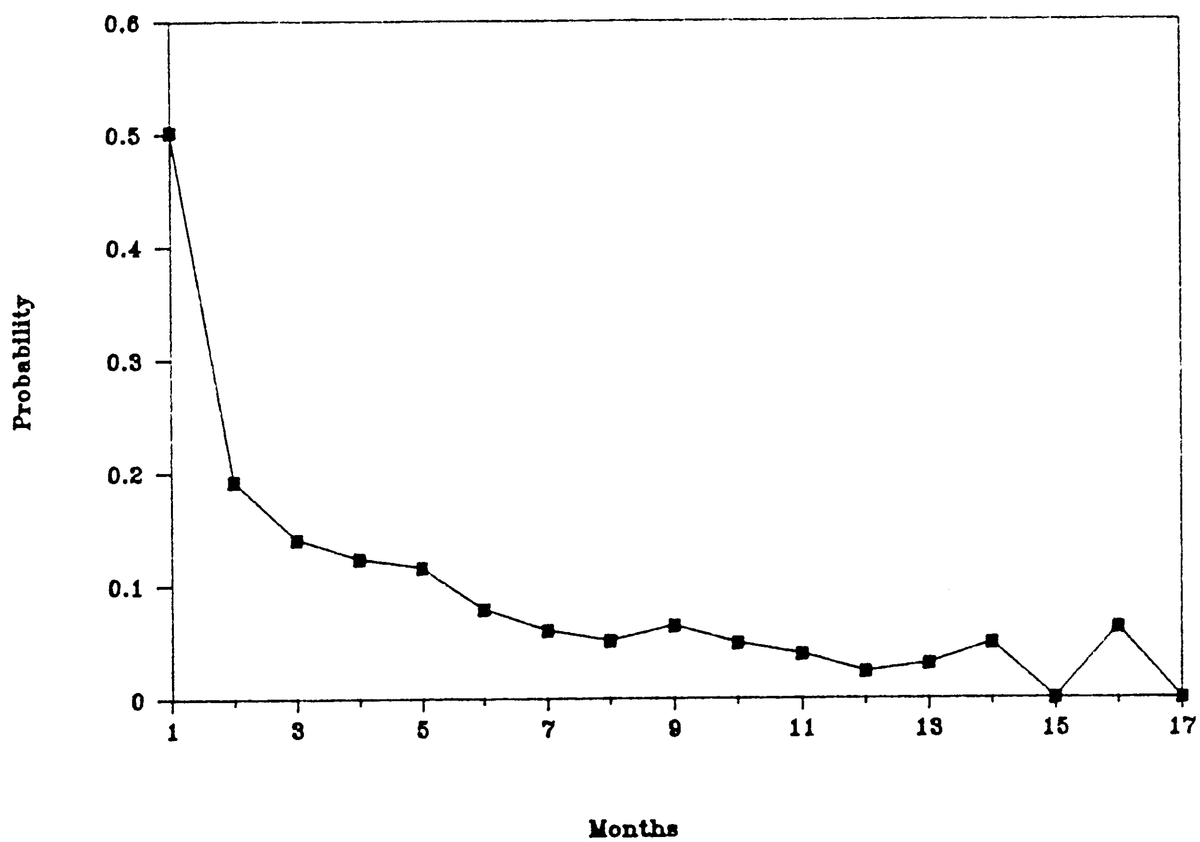
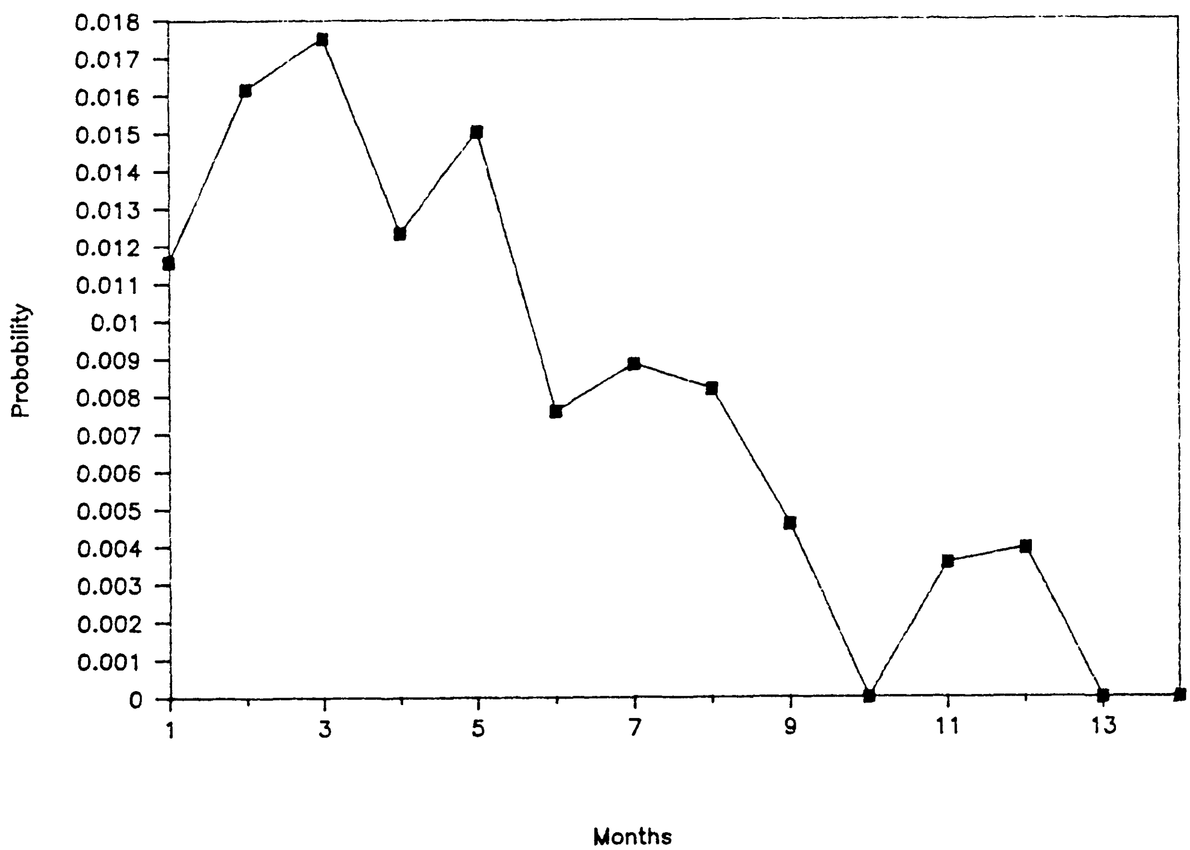


Figure 2 — New Job Hazard for Exit from Unemployment
Full 4,224 Spell Sample



the exhaustion of UI benefits. The former seems to be centered around 18 weeks. Conspicuously missing is the spike at 26 weeks.

The empirical hazard rate for new job finding is difficult to characterize. The hazard slopes upward for the first four months then declines througout the remaining months. Spikes are present at several periods, but given the relatively small sample sizes, should not be given too much weight. The downard slope is in contrast to Katz's PSID findings of an increasing hazard throughout and is inconsistent with the standard search theory models of unemployment. It may be that the differences in empirical hazards are the result of differences in sample selection criteria used in the two studies. Thus, differences in the hazard might result from differences in the observable characteristics of individuals in the two data sources. The hazard models estimated below are designed to account for this individual heterogeneity.

There are two possible interpretations of the absence of the 26 week spike. The first argues that the spike is merely an artifact of the discreteness of the data used in previous studies and that the continuous monitoring of the SIPP eliminates the lumpiness problem. This view seems unlikely in light of evidence from weekly administrative records suggesting that hazard rates increase in the weeks prior to benefit exhaustion (Meyer [1987]). A more plausible explanation is that these data cover a period of time in which Federal Supplemental Benefits extended eligibility to 39 weeks for most individuals. This explanation is consistent with the strong effect at the relevant months. There is evidence from the Continuous Wage and Benefit History for 1981 which indicates that potential UI benefits are considerably more likely to end at 39 weeks (Katz [1985]).

The presence of a spike at 18 weeks is considerably more difficult to explain. One possibility is that the 26 and 39 week exhaustion points do not apply to these individuals. In addition to limitations on the maximum number of weeks of benefits, a number of states impose ceilings on the maximum potential benefits paid. It is possible that these benefit ceilings yielded exhaustion of benefits at weeks prior to 26. While it is impossible, given the SIPP data, to determine when the actual exhaustion points are, it is interesting to note that of those states with potential benefits durations that vary with work history, a number have minimum potential durations in the neighborhood of 18 weeks.[10] Nevertheless, the magnitude of the 18 week spike is a puzzle.

In figures 3 and 4 I graph the corresponding sample hazards for the subsample. While there are differences between these and the full sample results, the two sets of hazards are mostly characterized

---

[10]For 1986, the only year for which data are available, 44 states have potential benefit durations which varied depending upon the individual work history. There were 9 states with minimum durations of 12 weeks, 3 with 13 weeks, 2 with 14 weeks, and 13 from 15-20 weeks.

Figure 3 — Recall Hazard for Exit from Unemployment

Subsample of 2,282 Spells

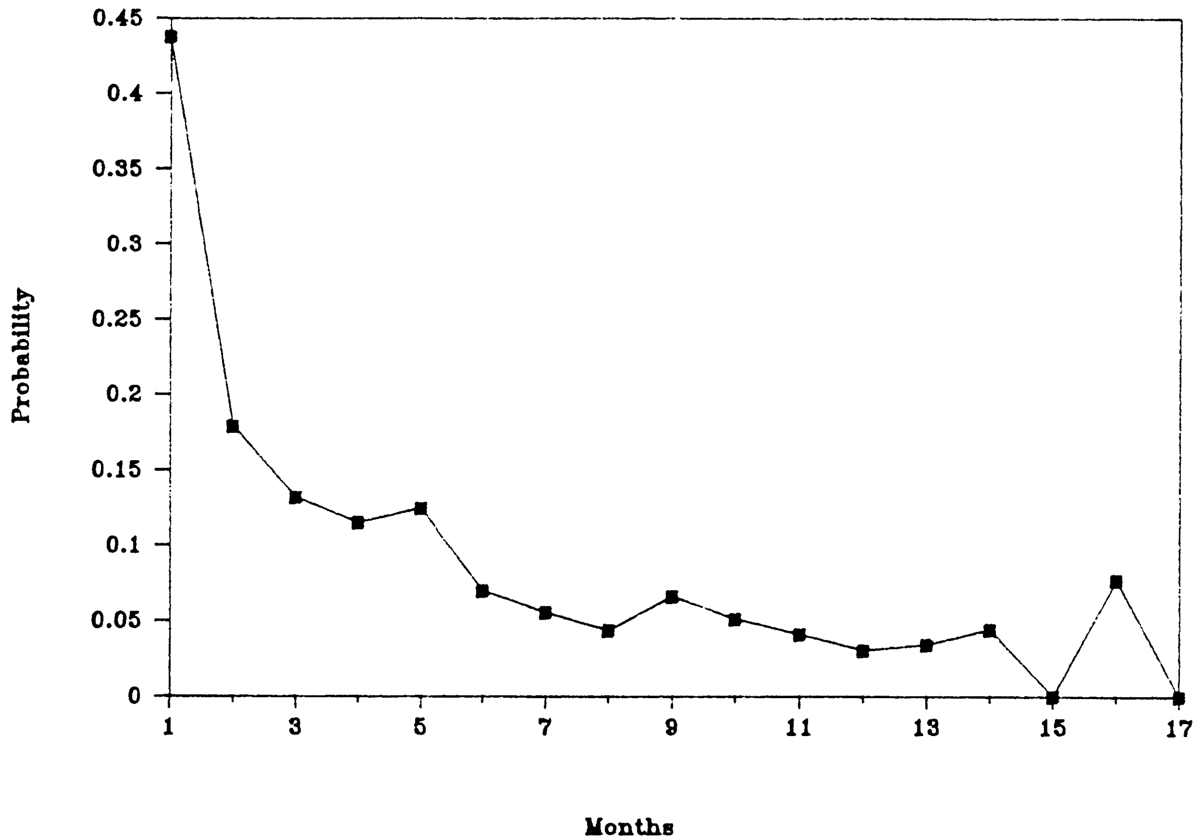
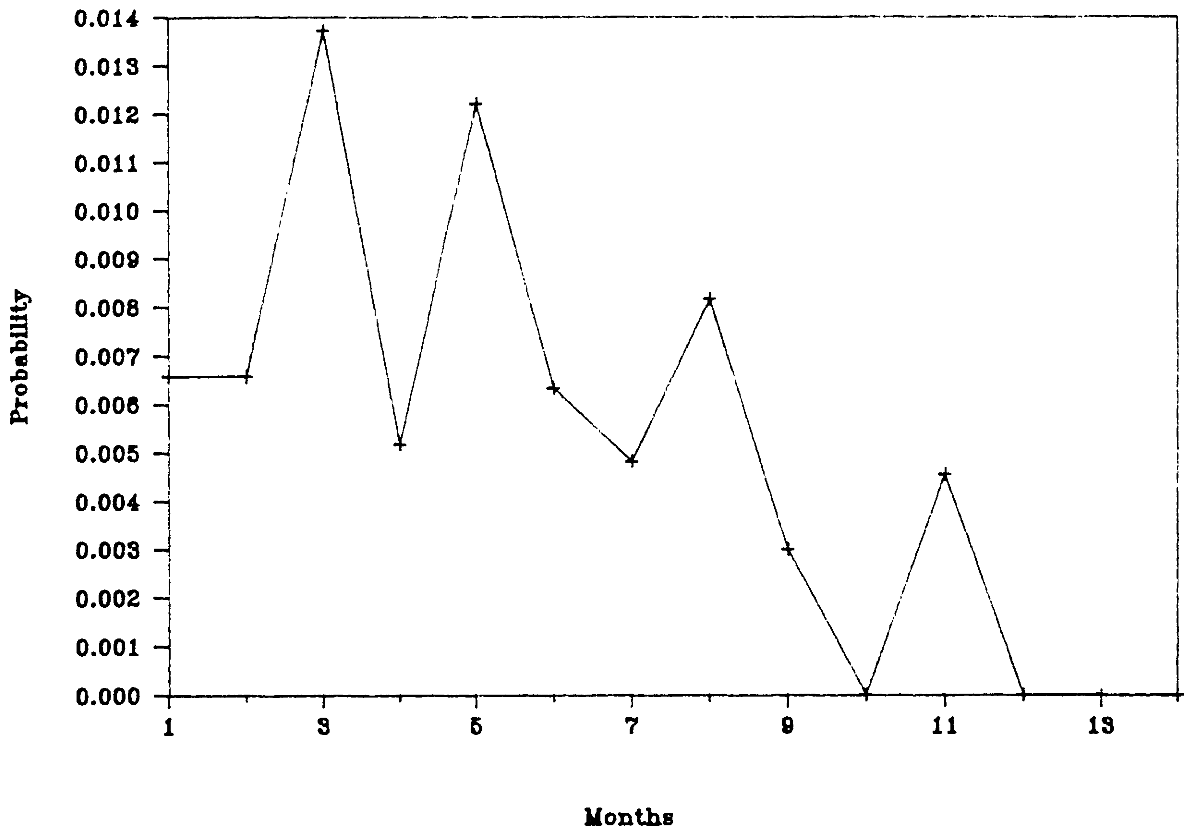


Figure 4 — New Job Hazard for Exit from Unemployment
Subsample of 2,282 Spells



by their similarities. In particular, the downward slope of both the recall and the latter portion of the new job hazard exists for the subsample. It is worth noting that the new job hazard exhibits more variability because of the very small sample size involved.

3.2 Single Risk Models

The models estimated in this chapter are based upon the Cox [1972] proportional hazards framework. Hazard, or duration, models have been used to study a variety of economic phenomena in which a single event occurs once over time. In essence, these models are reduced form specifications for conditional probabilities of an event occurring. Because of the obvious application to the study of unemployment durations, a number of researchers have, in recent years, applied these techniques to the study of unemployment spells.

The estimation of the hazard rates in this and the following section is based upon the semi-parametric techniques described in greater detail in Han and Hausman and in chapters 1 and 3 of this thesis. The actual models estimated were of the ordered logit form with the categories of the logit corresponding to the week of exit from unemployment or censoring. The model was estimated for 40 weeks because of relatively small numbers of exits beyond that point. Observations with durations greater than 40 weeks are censored at that point.

Estimation is performed using maximum likelihood techniques. I use a modified version of the algorithm described by Berndt, Hall, Hall and Hausman [1974]. Since the ordered logit form of the model implies global concavity of the likelihood function (Pratt [1981]), it is not surprising that convergence for the models was rapid, even for the models with large numbers of parameters.

Results for the single risk models are presented in tables 6 and 7. Table 6 contains parameter estimates for two models with basic heterogeneity controls. In the first model, the effect of UI is represented by an indicator for whether benefits are received; in the latter, a measure of the fraction of previous earnings that UI benefits represents, or the UI replacement rate, is used. The results are similar across the specifications. The coefficients for age, the non-white indicator, and the number of children under 18 in the family have precisely estimated negative coefficients, implying that individuals with those characteristics have relatively low exit hazards. Married individuals have higher hazards as do individuals with more education, though the coefficient for the latter variable is imprecisely measured. Income from assets reduces the hazard in both specifications, but the effect is not statistically different from zero.

For the basic specification, the extra information provided by incorporating benefit levels into the analysis appears to be important. The model using UI replacement rates fits the data better than the model using an indicator of benefit receipt, improving the

Table 6 - Single risk estimation for unemployment exit, basic heterogeneity controls models.

Heterogeneity Controls		
Age	-0.0163 (0.0013)	-0.0170 (0.0012)
Education	0.0005 (0.0065)	-0.0064 (0.0064)
Female (1=yes)	-0.0461 (0.0440)	-0.0574 (0.0434)
Married (1=yes)	0.0969 (0.0438)	0.1011 (0.0434)
Non-White (1=yes)	-0.1295 (0.0513)	-0.1327 (0.0512)
Number of Children Under 18	-0.0427 (0.0169)	-0.0410 (0.0168)
Income Variables		
Assets Income ($1000s per month)	-0.0569 (0.0729)	-0.0629 (0.0730)
UI Receipt Indicator (1=yes)	0.0134 (0.0454)	-----
UI Replacement Rate	-----	-0.4222 (0.0928)
Log Likelihood	-9803.22	-9788.87

Note: N = 4,224. Asymptotic standard errors in parentheses.

log likelihood by 15 from -9803.22 to -9788.87. Moreover, the coefficient for replacement rates has the expected sign, with benefits reducing the hazard, and is precisely measured. This result is in contrast to the positive, but insignificant coefficient for the receipt indicator.

To account for differences in spell durations resulting from variation across industries and occupations, I estimate a more general model incorporating dummy variables for industry and occupation class. The results are given in table 7. The base groups for the estimation are laborers and agriculture. The inclusion of classification variables improves the fit of the models substantially. Log-likelihood ratio tests for these models yield test statistic values of 84 and 96 respectively, both of which are statistically significant at the 5% level for 12 degrees of freedom. The occupation indicators seem to be the more important of the two groups of variables, with all of the occupation classes having low hazards relative to laborers. The lower hazards are surprising if one views unemployment as being a single spell phenomenon since one might think that, say, professionals would find jobs more readily available than laborers giving them a higher hazard. They are less puzzling if laborers are viewed as moving into and out of employment with some frequency so that over a given period the same aggregate unemployment duration is associated with a greater number of spells.

In contrast to the result in the basic model, the parameter for

Table 7 - Single risk estimation for unemployment exit hazards, full industry and occupation classification controls models.

Heterogeneity Controls			
Age	-0.0155 (0.0013)	-0.0161 (0.0013)	-0.0170 (0.0013)
Education	0.0188 (0.0072)	0.0177 (0.0071)	0.0192 (0.0071)
Female (1=yes)	0.0745 (0.0486)	0.0760 (0.0484)	0.0507 (0.0483)
Married (1=yes)	0.0987 (0.0438)	0.1015 (0.0435)	0.1076 (0.0435)
Non-White (1=yes)	-0.1307 (0.0522)	-0.1334 (0.0520)	-0.1332 (0.0516)
Number of Children Under 18	-0.0481 (0.0170)	-0.0476 (0.0169)	-0.0427 (0.0167)
Income Variables			
Assets Income ($1000s per month)	-0.0368 (0.0762)	-0.0419 (0.0761)	-0.0316 (0.0749)
UI Receipt Indicator (1=yes)	-0.0556 (0.0467)	------	------
UI Replacement Rate	------	-0.4963 (0.0925)	------
UI Benefit Level (100$ per week)	------	------	-0.2248 (0.0266)

Table 7 (continued)

Occupation Classification			
Professional	-0.4041	-0.4148	-0.4302
	(0.0777)	(0.0774)	(0.0764)
Sales	-0.2561	-0.2610	-0.2809
	(0.0630)	(0.0629)	(0.0617)
Service	-0.2336	-0.2403	-0.2618
	(0.0688)	(0.0682)	(0.0669)
Craft	-0.0398	-0.0538	-0.0597
	(0.0528)	(0.0530)	(0.0530)
Industry Classification			
Mining	-0.1765	-0.1329	-0.0312
	(0.2672)	(0.2609)	(0.2556)
Construction	0.1759	0.2072	0.2780
	(0.1909)	(0.1897)	(0.1886)
Durables	-0.0482	-0.0371	-0.0032
	(0.1910)	(0.1898)	(0.1886)
Non-durables	0.1161	0.1274	0.2049
	(0.1881)	(0.1868)	(0.1856)
Transportation	-0.0356	-0.0234	0.0081
	(0.1930)	(0.1920)	(0.1910)
Wholesale	-0.0570	-0.0452	-0.0417
Durables	(0.2288)	(0.2276)	(0.2256)
Wholesale	-0.1190	-0.1190	-0.1067
Non-durables	(0.1878)	(0.1866)	(0.1856)
Service	0.0844	0.0869	0.1057
Industries	(0.1835)	(0.1824)	(0.1814)
Log Likelihood	-9761.14	-9742.70	-9696.65

Note: N = 4,224. Asymptotic standard errors in parentheses.

education is now precisely measured, with additional years of schooling increasing the hazard rate. It is not surprising that controlling for occupation and industry has an effect on the education variable since the correlation between education and industry and occupation is usually quite high. The coefficient for UI in the first specification switches sign so that benefit receipt is associated with lower hazards, but is still statistically insignificant. A higher replacement rate continues to reduce exit probabilities.

Because of some concern over the accuracy of the replacement rate variable, I specify a third model that incorporates directly UI benefit levels. The results are presented in the third column of tables 6 and the second column of table 7. Interestingly, the model fits the data slightly better, though the results for the other variables are similar to the results obtained when the replacement rate is employed.

Since the sample of 4,224 observations represents an unbalanced panel with multiple observations on individuals, the presence of individual specific random effects would yield inconsistent parameter estimates. As a crude indicator of the sensitivity of the model to the inclusion of multiple spells, I reestimated the models given in tables 6 and 7 for the subsample containing one observation for each individual. The results of estimation for this sample are presented in tables 8 and 9.

Table 8 - Single risk estimation for unemployment exit hazards, full industry and occupation classification controls models. Subsample of spells.

Heterogeneity Controls		
Age	-0.0161 (0.0017)	-0.0165 (0.0017)
Education	-0.0058 (0.0085)	-0.0074 (0.0084)
Female (1=yes)	-0.0727 (0.0599)	-0.0755 (0.0592)
Married (1=yes)	0.0694 (0.0592)	0.0670 (0.0590)
Non-White	-0.2412 (0.0682)	-0.2421 (0.0680)
Number of Children under 18	-0.0498 (0.0226)	-0.0497 (0.0225)
Income Variables		
Asset Income ($1000s per month)	-0.1192 (0.0994)	-0.1225 (0.0989)
UI Replacement Rate	-0.0417 (0.0617)	-----
UI Benefits ($100 per week)	-----	-0.3100 (0.1135)
Log Likelihood	-5349.18	-5344.70

Note: N = 2,282. Asymptotic standard errors in parentheses.

Table 9 - Single risk estimation for unemployment exit hazards, full industry and occupation classification controls models. Subsample of spells.

Heterogeneity Controls		
Age	-0.0148 (0.0018)	-0.0151 (0.0018)
Education	0.0102 (0.0097)	0.0092 (0.0097)
Female	0.0457 (0.0679)	0.0480 (0.0676)
Married	0.0934 (0.0595)	0.0896 (0.0593)
Non-White	-0.2112 (0.0693)	-0.2118 (0.0691)
Number of Children under 18	-0.0511 (0.0227)	-0.0523 (0.0226)
Income Variables		
Asset Income ($1000 per month)	-0.0951 (0.1047)	-0.0987 (0.1045)
UI Replacement Rate	-0.1134 (0.0639)	-----
UI Benefits ($100 per week)	-----	-0.3680 (0.1140)

Table 9 - (continued)

Occupation Classification		
Professional	-0.3288 (0.1064)	-0.3268 (0.1061)
Sales	-0.1425 (0.0882)	-0.1392 (0.0884)
Service	-0.1553 (0.0962)	-0.1448 (0.0959)
Craft	0.0126 (0.0740)	0.0108 (0.0740)
Industry Classification		
Mining	-0.1282 (0.3510)	-0.0729 (0.3452)
Construction	0.2910 (0.2661)	0.3038 (0.2669)
Durables	-0.1666 (0.2680)	-0.1508 (0.2690)
Non-durables	0.1568 (0.2615)	0.1709 (0.2625)
Transportation	-0.0977 (0.2695)	-0.0743 (0.2706)
Wholesale Durables	-0.1921 (0.3293)	-0.1873 (0.3297)
Wholesale Non-durables	-0.0590 (0.2603)	-0.0518 (0.2613)
Service	0.0952 (0.2532)	0.1032 (0.2542)
Log Likelihood	-5322.29	-5317.73

Note: N = 2,282. Asymptotic standard errors in parentheses.

The results for the subsample are quite similar to those for the full sample. Comparing, for example the results for the benefit specification in column 3 of tables 7 and 9, only minor differences are observed in the heterogeneity control variables. The influence of UI benefits is somewhat larger in the subsample (0.3680 versus 0.2248), but so too is the standard error. Similarly, the industry and occupation results differ by amounts that are within the realm of sampling error. While this does not rigorously prove that individual effects are not likely to be a problem, it does suggest that the conclusions are unlikely to be altered by considering the more general model.

Taken together, the results from the single risk models indicate that UI receipt has an important effect upon unemployment exit rates. UI is also important relative to demographic characteristics. For example, in the full industry and occupation control model with UI replacement rates (column 2 of table 7), a change in the rate of ten basis points, from say .40 to .50, has an impact almost four times as large as the effect of being non-white. Similarly, the indicator for receipt of benefits (column 1 of table 7) has about the same impact as two to three fewer years of education, while an extra $10 a week in UI benefits has is equivalent to being two years older in reducing hazards. Furthermore, $10 additional dollars of weekly UI benefits has about the same impact as receiving an additional $1000 in assets income per month, suggesting that the direct relationship between employment status and the receipt of benefits is important.

To get a better sense of the impact that these values have on the hazard rates, I calculate elasticities of the hazard at the mean values for the data. Given the proportional hazard specification, the elasticity takes a particularly convenient form. Recalling that the hazard function is of the form $\lambda(t|X,\beta) = \lambda_o(t)\ exp(X\beta)$, it follows that

$$\eta_k = \frac{\partial\lambda(t|X,\beta)}{\partial X_k} \ \frac{X_k}{\lambda(t|X,\beta)} = \frac{\lambda_o(t)\ exp(X\beta)\ X_k\beta_k}{\lambda_o(t)\ exp(X\beta)} = X_k\beta_k.$$

The elasticities corresponding to the models in table 7 are given in table 10. The elasticities for UI replacement rate and benefit levels are relatively significant, exceeding in magnitude all variables with the exception of age. According to these figures, an increase in the benefit level of 10 percent per week would uniformly increase the hazard rate by about 2.3 percent.

3.3 Competing Risks Models

To account for the the fact that individuals may exit from unemployment by finding a new job or by being recalled to their previous one, I estimate competing risks models of unemployment duration. The basic framework is set forth elsewhere in this thesis. Since, under the assumption that the hazard rates conditional upon

Table 10 - Elasticities of covariates evaluated at means for single risk model with full industry and occupation controls. Full sample of spells.

Heterogeneity Controls			
Age	0.6473 (0.0543)	0.6723 (0.0543)	0.7099 (0.0543)
Education	-0.2277 (0.0872)	-0.2143 (0.0860)	-0.2325 (0.0860)
Female (1=yes)	-0.0246 (0.0160)	-0.0251 (0.0160)	-0.0167 (0.0159)
Married (1=yes)	-0.0582 (0.0258)	-0.0599 (0.0257)	-0.0635 (0.0257)
Non-White (1=yes)	0.0183 (0.0073)	0.0187 (0.0073)	0.0186 (0.0072)
Number of Children Under 18	0.0409 (0.0145)	0.0405 (0.0144)	0.0363 (0.0142)
Income Variables			
Assets Income ($1000s per month)	0.0025 (0.0052)	0.0029 (0.0052)	0.0022 (0.0052)
UI Receipt (1=yes)	0.0122 (0.0103)	-----	-----
UI Replacement Rate	-----	0.2652 (0.0314)	-----
UI Benefits ($100s per week)	-----	-----	0.2343 (0.0437)

Note: The mean values for replacement rate and benefits are conditional upon receipt. Mean values for the other covariates are taken from the first column of table 2.

individual characteristics are independent, the likelihood function factors into separate components for each risk (Kalbfleisch and Prentice [1980]), the analysis involves maximum likelihood estimation of two separate ordered logit models in which first one and then the other of the failure types is treated as a censored failure. Competing risks models are not estimated for the individual subsample due to the exceedingly small cell sizes associated with the new job failure type. Preliminary attempts at estimating a competing risks model with a general error structure failed to achieve convergence.

The results for two competing risks specifications are presented in table 11. Furthermore, the model using UI benefit levels fits the data better than the corresponding replacement rate model, with an improvement in the log likelihood of the sample from -9360.72 to -9318.26 for the recall hazard and from -903.755 to -901.601 for the new job hazard. Once again, the results are essentially the same whether UI is represented by the replacement rate or by benefit levels.

Differences between the coefficients across risks illustrate the need to account for the competing risks nature of the exit from unemployment. The most significant differences between the relative effects of covariates are for the heterogeneity controls. Age, education, and non-white all have larger relative effects upon the new job hazard than upon the recall hazard; the age and education coefficients are three times as large as, and the coefficient for the

Table 11 - Competing risks estimation for unemployment exit hazards, full industry and occupation classification controls models. Full sample of spells.

Variables	Recall	New Job	Recall	New Job
Heterogeneity Controls				
Age	-0.0146 (0.0013)	-0.0464 (0.0092)	-0.0155 (0.0013)	-0.0474 (0.0092)
Education	0.0143 (0.0073)	0.1044 (0.0396)	0.0159 (0.0072)	0.1000 (0.0397)
Female (1=yes)	0.0988 (0.0498)	-0.3570 (0.2452)	0.0731 (0.0497)	-0.3822 (0.2462)
Married (1=yes)	0.1213 (0.0446)	-0.2201 (0.2222)	0.1275 (0.0446)	-0.1990 (0.2223)
Non-White (1=yes)	-0.0962 (0.0528)	-1.1129 (0.4212)	-0.0955 (0.0525)	-1.1042 (0.4206)
Number of Children Under 18	-0.0494 (0.0170)	0.0423 (0.0916)	-0.0448 (0.0168)	0.0402 (0.0917)
Income Variables				
Assets Income ($1000s per month)	-0.0291 (0.0779)	-0.5348 (0.7355)	-0.0197 (0.0765)	-0.5094 (0.7491)
UI Replacement Rate	-0.5249 (0.0929)	-0.0605 (0.3409)	-----	-----
UI Benefit Level (100$ per week)	-----	------	-0.2279 (0.0270)	-0.1720 (0.1254)

Table 11 - (continued)

Occupation Classification				
Professional	-0.4135	-0.4496	-0.4280	-0.4586
	(0.0789)	(0.4289)	(0.0777)	(0.4284)
Sales	-0.2865	0.2640	-0.3049	0.2441
	(0.0644)	(0.3496)	(0.0632)	(0.3486)
Service	-0.2765	0.3779	-0.2966	0.3566
	(0.0702)	(0.3983)	(0.0689)	(0.3982)
Craft	-0.0731	0.3894	-0.0780	0.3772
	(0.0540)	(0.3296)	(0.0539)	(0.3292)
Industry Classification				
Mining	-0.1721	0.5520	-0.0814	0.6710
	(0.2718)	(1.2064)	(0.2657)	(1.2004)
Construction	0.1929	0.5207	0.2622	0.6421
	(0.1935)	(1.0893)	(0.1920)	(1.0914)
Durables	-0.0485	0.1900	-0.0134	0.2234
	(0.1936)	(1.0824)	(0.1921)	(1.0837)
Non-durables	0.1254	0.2356	0.2016	0.3327
	(0.1903)	(1.0711)	(0.1889)	(1.0734)
Transportation	-0.0323	0.2627	-0.0011	0.3194
	(0.1956)	(1.1009)	(0.1943)	(1.1028)
Wholesale	-0.0396	-0.1790	-0.0311	-0.1565
Durables	(0.2334)	(1.3006)	(0.2312)	(1.2991)
Wholesale	-0.1449	0.4336	-0.1321	0.4553
Non-durables	(0.1904)	(1.0812)	(0.1891)	(1.0828)
Service	0.0763	0.4180	0.0959	0.4494
	(0.1864)	(1.0835)	(0.1851)	(1.0852)
Log Likelihood	-9360.72	-903.755	-9318.26	-901.601

Note: Asymptotic standard errors in parentheses.

non-white indicator is ten times greater than the corresponding coefficients for the recall hazard. The sensitivity of the new job hazard is consistent with the view that observable individual characteristics are more important for an individual in finding a new job than in being recalled. A similar result is observed for being female, but the coefficients are not precisely estimated.

The coefficient for UI replacement is large and statistically significant for recall but small and insignificant for the new job risk. The coefficients for UI benefit levels are similar in character. The finding that UI affects recall more than new job finding is counterintuitive, since if anything, one would expect UI to have a larger impact upon new job hazards than upon the probability of being recalled. It may be that the small sample size for new job individuals makes it difficult to estimate the impact of benefit level variables with any precision. The small standard errors with which the heterogeneity coefficients are estimated, however, makes it difficult to support this argument.

Even if this is the case, the strong relationship between UI and recall suggests that firms are taking advantage of the subsidy to layoff provided by UI systems. If there were no firm response to the UI system, then there should be no observable impact upon recall hazards. If firm layoff and recall behavior are, in fact, affected by UI, then standard models of unemployment duration are missing a large part of the story underlying the effect of UI on unemployment spells.

There is a strong occupational component to the recall risk. Each of the occupation classes exhibits a smaller risk of exit than laborers with all but the craft classification statistically significant. The largest differential is observed for professional workers. In contrast, the occupational dummies for new job are estimated with considerable error. The industry coefficients all have relatively large standard errors.

The hazard estimates provide mixed support for the search interpretation of unemployment durations. The baseline hazards for the competing risks models are presented in tables 12 and 13 along with the associated asymptotic standard errors. Weekly hazards for the model using UI benefits are also depicted in figures 5 and 6. The baseline hazard for recall is very similar to the empirical hazard depicted earlier, with small spikes present at the same places. More significantly, upon correcting for the presence of observable heterogeneity and estimating the competing risk model, the baseline hazard rate for new job finding appears to be upward sloping, at least for the first few months. There are spikes in the hazard just prior to 18, 26 and 39 weeks, a result that is consistent with the view that UI exhaustion matters for job search.

The new job hazard for monthly exit times is presented in figure 7 for comparison with the weekly hazards and with the monthly empirical hazard depicted earlier. The increase in the hazard for

Table 12 - Estimated weekly baseline hazards for dual risk model of unemployment spell duration, replacement rate specification.

	Recall		*New Job*	
Week	Hazard	Std Err	Hazard	Std Err
1	0.3981	0.0683	0.0043	0.0054
2	0.2810	0.0537	0.0036	0.0047
3	0.1905	0.0393	0.0051	0.0065
4	0.1577	0.0337	0.0090	0.0106
5	0.1171	0.0261	0.0075	0.0095
6	0.0919	0.0215	0.0053	0.0068
7	0.0914	0.0215	0.0058	0.0075
8	0.0728	0.0180	0.0055	0.0069
9	0.0811	0.0198	0.0069	0.0085
10	0.0712	0.0179	0.0053	0.0071
11	0.0615	0.0161	0.0080	0.0102
12	0.0608	0.0161	0.0074	0.0095
13	0.0738	0.0190	0.0065	0.0087
14	0.0589	0.0161	0.0014	0.0023
15	0.0871	0.0220	0.0090	0.0121
16	0.0238	0.0085	0.0034	0.0049
17	0.1127	0.0275	0.0104	0.0131
18	0.0696	0.0194	0.0099	0.0131
19	0.0137	0.0068	0.0023	0.0037
20	0.0445	0.0147	0.0048	0.0070
21	0.0318	0.0117	0.0025	0.0045
22	0.0594	0.0182	0.0051	0.0073
23	0.0415	0.0145	0.0056	0.0080
24	0.0266	0.0110	0.0000	0.0000
25	0.0239	0.0104	0.0058	0.0082
26	0.0457	0.0159	0.0032	0.0050
27	0.0186	0.0092	0.0034	0.0056
28	0.0342	0.0135	0.0034	0.0056
29	0.0320	0.0132	0.0074	0.0103
30	0.0167	0.0091	0.0000	0.0000
31	0.0212	0.0105	0.0038	0.0061
32	0.0351	0.0146	0.0041	0.0066
33	0.0092	0.0068	0.0043	0.0070
34	0.0554	0.0196	0.0044	0.0069
35	0.0583	0.0207	0.0000	0.0000
36	0.0119	0.0088	0.0000	0.0000
37	0.0244	0.0133	0.0000	0.0000
38	0.0063	0.0065	0.0000	0.0000
39	0.0323	0.0160	0.0000	0.0000
40	0.0412	0.0190	0.0000	0.0000

Table 13 - Estimated weekly baseline hazards for dual risk model of unemployment spell duration, UI benefit specification.

Week	*Recall* Hazard	*Recall* Std Err	*New Job* Hazard	*New Job* Std Err
1	0.4067	0.0688	0.0047	0.0059
2	0.2904	0.0547	0.0040	0.0052
3	0.1987	0.0405	0.0057	0.0073
4	0.1656	0.0350	0.0101	0.0120
5	0.1236	0.0273	0.0086	0.0109
6	0.0972	0.0226	0.0061	0.0078
7	0.0968	0.0226	0.0067	0.0087
8	0.0772	0.0189	0.0063	0.0080
9	0.0862	0.0209	0.0079	0.0099
10	0.0758	0.0189	0.0061	0.0082
11	0.0655	0.0170	0.0092	0.0118
12	0.0648	0.0170	0.0085	0.0109
13	0.0786	0.0201	0.0076	0.0101
14	0.0627	0.0170	0.0016	0.0027
15	0.0927	0.0232	0.0105	0.0141
16	0.0252	0.0090	0.0038	0.0056
17	0.1198	0.0290	0.0119	0.0149
18	0.0743	0.0205	0.0116	0.0152
19	0.0147	0.0073	0.0027	0.0043
20	0.0476	0.0155	0.0055	0.0081
21	0.0338	0.0125	0.0028	0.0052
22	0.0633	0.0193	0.0058	0.0084
23	0.0441	0.0153	0.0063	0.0091
24	0.0283	0.0117	0.0000	0.0000
25	0.0254	0.0110	0.0067	0.0094
26	0.0487	0.0168	0.0037	0.0057
27	0.0198	0.0098	0.0039	0.0063
28	0.0364	0.0143	0.0040	0.0066
29	0.0338	0.0139	0.0084	0.0118
30	0.0177	0.0096	0.0000	0.0000
31	0.0225	0.0112	0.0044	0.0069
32	0.0372	0.0153	0.0047	0.0075
33	0.0096	0.0070	0.0049	0.0080
34	0.0586	0.0206	0.0050	0.0080
35	0.0616	0.0218	0.0000	0.0000
36	0.0125	0.0093	0.0000	0.0000
37	0.0260	0.0141	0.0000	0.0000
38	0.0066	0.0068	0.0000	0.0000
39	0.0339	0.0169	0.0000	0.0000
40	0.0435	0.0200	0.0000	0.0000

Figure 5 – Estimated Weekly Recall Hazard

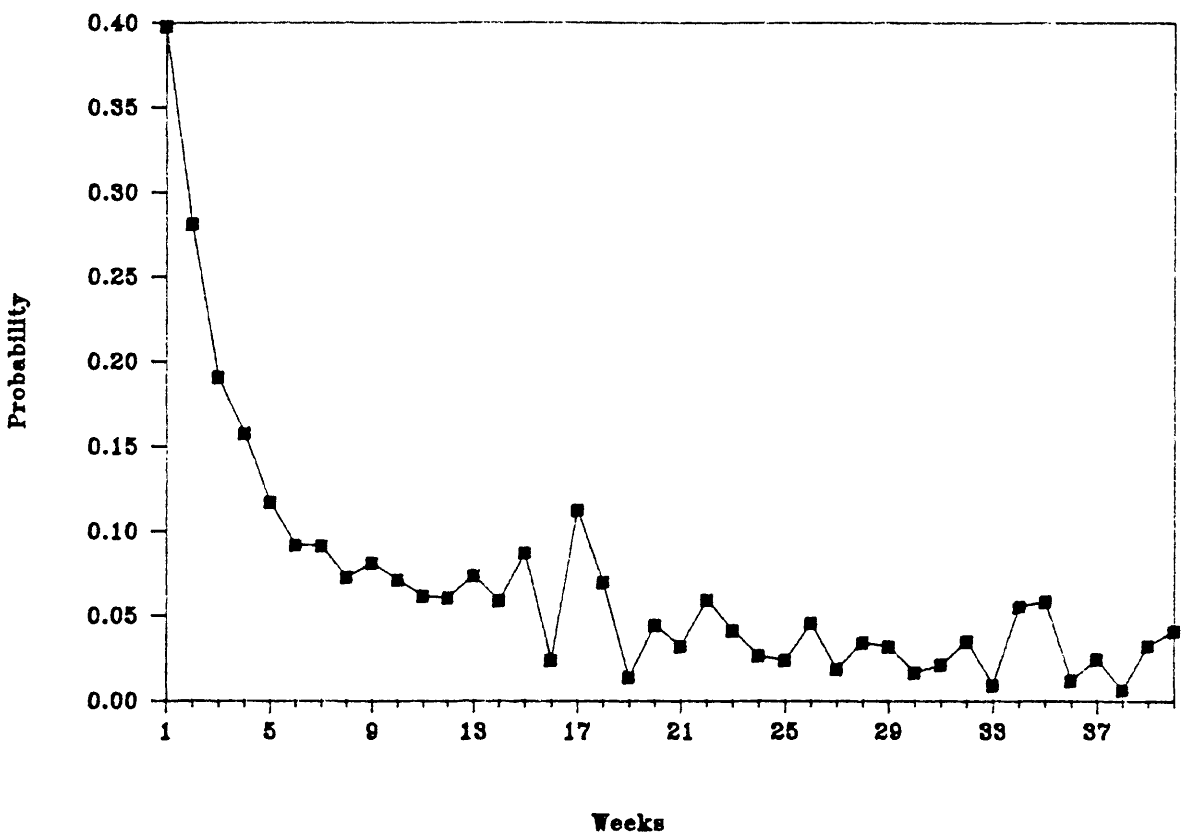
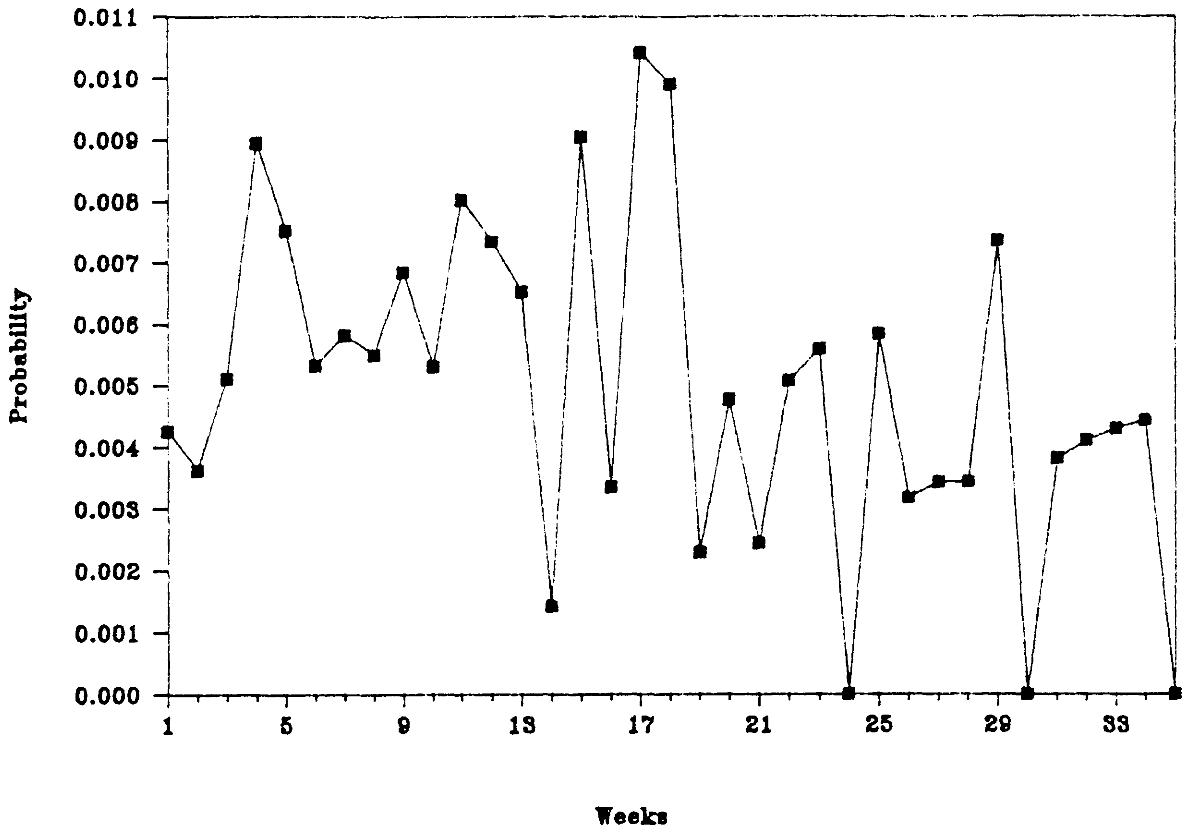


Figure 6 — Estimated Weekly New Job Hazard



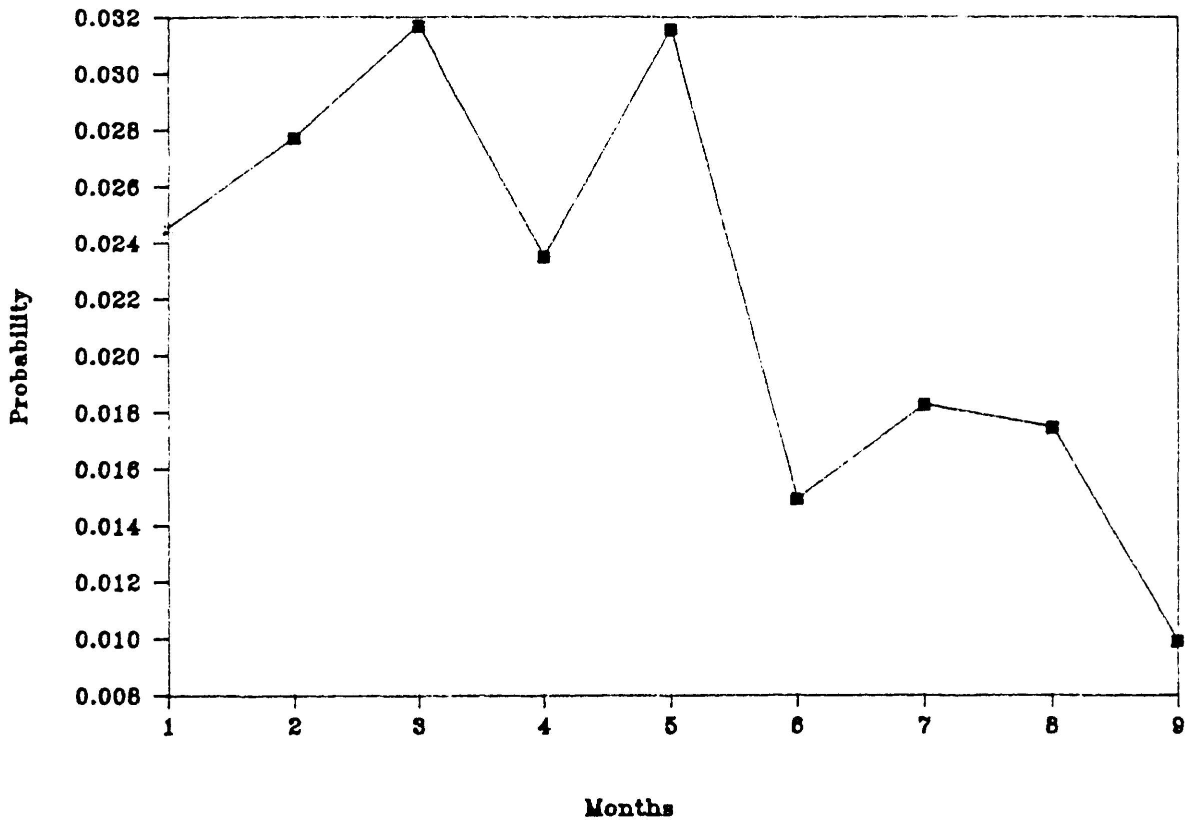
the first few months is again apparent, as is the sharp decline in the subsequent months.

One really should be cautioned against drawing strong conclusions from the new job hazard shapes since, given the large standard errors in tables 12 and 13, figures 6 and 7 are consistent with virtually any shape. Nevertheless based upon the point estimates of the baseline hazard, the evidence in support of the standard search model is at best mixed, with the increasing hazard for the first five months weakly supporting the search interpretation, and the subsequent non-monotonicity at variance with the predictions of the model.

The evidence of a downward sloping portion of the new job hazard is contrary to Katz's findings using PSID data. Some possible explanations for the differences between the two results have already been suggested, namely the presence of heterogeneity bias or the fact that the basic composition of the two samples differs substantively. The latter explanation, which relies on the existence of very different layoff and search behavior across industry and occupation types, implies that simple models of unemployment and search are inadequate in explaining behavior.

Finally, the one unexplained finding in the estimates of the models has been the spike in the hazard rates for both new job and recall at around 18 weeks. There is no obvious explanation for the importance of this spike, though the spike may very well be related

Figure 7 — Estimated Monthly New Job Hazard



to the exhaustion of UI benefits. Additional work will be required to determine whether it is merely an artifact of the data and the techniques used for calculating spells in the SIPP data, or whether there is a more fundamental relationship that has been missed.

6. *Conclusion*

This analysis of unemployment durations using the SIPP data takes advantage of the benefits accruing to the use of this data. In particular, I am able to examine simultaneously the effects of UI benefit receipt on the hazards of exit from unemployment via either new job or recall. The durations used in this analysis are calculated on the basis of weekly data on labor force status and are therefore not subject to much of the error in calculating durations that has resulted from the use of alternative data sources.

At the same time, the current analysis does not take full advantage of a great wealth of previously unavailable information provided by the SIPP. In particular, the future availability of the work history supplement should allow researchers to examine questions about labor force mobility in considerably greater detail. Moreover, the supplement will allow for the verification of some of the more curious aspects of labor force behavior found in this and other papers. For those reasons, the results presented here should be considered a preliminary examination of the SIPP data.

Having said that, there are some important lessons that can be taken away from the current analysis. First, based upon the spells of unemployment calcualted from the weekly SIPP data, it appears that the current emphasis on examining the duration of a given spell of unemployment is missplaced. The very large numbers of short term spells combined with the large fraction of individuals with multiple spells points to a need to examine the question of duration in conjunction with the question of spell frequency. This is not a new idea, but it has been relatively neglected in recent empirical work.

The results of this paper also suggest that while UI has the expected positive impact upon the duration of a spell, the relationship is more complex than as described by search theory. For one thing, the strong influence of UI in altering the recall hazard suggests that the interaction between UI and firm layoff policy deserves additional empirical attention.

Finally, given the differences in behavior across risks implied by the estimates of coefficients and the baseline hazard in the competing risk models, the important task of identifying the differences between behavior associated with recall and with new job finding first undertaken by Katz should be continued. Despite the mixed support for the standard search models in the SIPP data, it may well be that examining the duration of unemployment spells in a more general framework will yield results that support the implications of a standard search model. But this preliminary analysis suggests that steps in the direction of a more general framework are still needed.

Chapter 3

Semi-parametric Proportional Hazards Estimation of Single and Competing Risks Models with Time-varying Covariates

Introduction

Economists increasingly have become involved in the development and use of techniques for the analysis of duration data. Duration models provide a convenient framework for the examination of economic phenomena which yield data on the length of time before an event occurs. The length of unemployment spells, the number of years to retirement, and the duration of a labor strike are but a few of the applications to which hazard models have been applied.

Since its introduction by Cox [1972], the proportional hazards model has become the standard for the examination of failure time data. A number of refinements of the basic Cox model have been developed. Recent econometric research on proportional hazards duration models has focused on estimation methods which reduce the number of functional form restrictions. This effort has been motivated by concern over possible biases resulting from parametric misspecification of either the baseline hazard or the distribution of unobserved heterogeneity.

While techniques for the non-parameteric estimation of the heterogeneity distribution have been proposed, these generally have been carried out in the context of a parametric specification for the baseline hazard (see, for example, Heckman and Singer [1985]). Work by Han and Hausman [1986] and Manton, Stallard and Vaupel [1986] suggests that the biases in the proportional hazards framework may be larger for misspecification of the baseline hazard than the for misspecified heterogeneity distributions. Accordingly, it should be more productive to consider the relaxation of restrictions on the baseline hazard.

There exist several techniques for the estimation of models without parametric restrictions upon the baseline hazard, but each of these has some limitations on general applicability. The Kaplan-Meier [1958] product limit estimator is limited by the assumption of a homogeneous population. Cox's [1972] partial likelihood technique handles naturally the inclusion of covariates, but the introduction of unobservable heterogeneity is computationally intractable. Han and Hausman [1986] propose an estimation technique that is based upon the regression form of the proportional hazards model and allows for covariates, heterogeneity and competing risks. The first two chapters of this thesis consist of applications of these techniques to the study of retirement and unemployment. However, the properties of this estimator are only shown to hold for the special case where the covariates are constant over time. A related technique has been suggested by Prentice and Gloeckler [1978] which allows for time-varying covariates, but has not been appplied to the case of competing risks. Given the importance of panel data in economic research and the number of cases in which there is more than one risk, a model that utilizes these semi-parametric estimation techniques but allows for time-varying covariates and competing modes of failure is needed.

In this chapter, I consider a more general specification of the regression form of the proportional hazards likelihood function used in chapters 1 and 2. The more general form of the model admits time-varying covariates and extends readily to the competing risks framework. This specification is an extension of the Han and Hausman technique and posesses the natural advantages of that estimator without the limitations upon the distribution of covariates over time. I also demonstrate identification and asymptotic normality of both the single and dual risk estimators.[1]

In section 1, I outline the specification of the semi-parameteric single risk hazard model and demonstrate identification and asymptotic normality of the estimator. Section 2 contains analagous results for the competing risk case. There is a brief concluding section.

---

[1]The single risk model considered here is merely a transformation of the Prentice and Gloeckler model to the regression form so that the identification results are not particularly original. They are presented along with the original dual risk results in an effort to provide a unified framework for the estimation of semi-parametric models with time-varying covariates.

1. Single-Risk Duration Models

1.1. Specification

Following standard hazard specification practice the hazard rate for individual i is assumed to be of the proportional hazards (Cox [1972]) form:

$$(1.1) \qquad \lambda(t|X_i(t),\beta) = \lambda_o(t)\ exp(X_i(t)\beta)$$

where $\lambda_o(t)$ is a possibly unknown function of the elapsed duration and $X_i(t)'$ and β are $k \times 1$ column vectors. This implies that the probability of observing a (continuous) failure time at time period t is given by (suppressing the subscript for individual i):

$$(1.2) \qquad f(t|X,\beta) = \lambda_o(t)\ exp(X(t)\beta)\ exp\ (-\int_0^t \lambda_o(s)\ exp(X(s)\beta)\ ds)$$

In the special case where the covariates do not vary over time so that $X(t) = X \ \forall\ t$, this specification can be rewritten in the equivalent regression form:

$$(1.3) \qquad f(t|X,\beta) = Pr(log \int_0^t \lambda_o(s)\ ds + X\beta = \epsilon)$$

$$= \qquad Pr(\ell_t + X\beta = \epsilon)$$

where ϵ has an extreme value distribution and $\ell_t \equiv \int_0^t \lambda_o(s)\ ds$. Han and Hausman (HH) [1986] show that this specification implies that the hazard model may be estimated without parametric specification of the baseline hazard function $\lambda_o(t)$. The estimation technique involves maximizing an ordered logit likelihood where the categories are defined by the failure times. The constant terms in the model correspond to functions of the baseline hazard evaluated at discrete points, so that the unknown function $\lambda_o(t)$ is estimated as a step function.

The extension of this model to the case of time-varying covariates is relatively straightforward, and begins with the standard change of variables to derive the alternative form of the likelihood function. First, recall that the likelihood of observing a failure time at time period t is given by (1.2). Define the transformation:

$$\Upsilon(t) = \int_0^t \lambda_o(s)\ exp(X(s)\beta)\ ds \tag{1.4}$$

$$d\Upsilon = \lambda_o(t)\ exp(X(t)\beta)\ dt$$

$$\frac{dt}{d\Upsilon} = \left[\lambda_o(t)\ exp(X(t)\beta)\right]^{-1}$$

and consider the resulting change of variables from t to Υ.

$$(1.5) \qquad f_{\varphi}(\varphi|X,\beta) = f_{t}(\varphi^{-1}(t)\,|X,\beta) \left|\frac{dt}{d\varphi}\right| = exp(-\varphi)$$

which is the density function of an exponential random variable.

Define a second transformation from φ to ϵ:

$$(1.6) \qquad \begin{aligned} \epsilon(\varphi) &= log(\varphi) \\ d\epsilon &= \frac{1}{\varphi}\,d\varphi \\ \frac{d\varphi}{d\epsilon} &= \varphi \end{aligned}$$

so that

$$(1.7) \qquad f_{\epsilon}(\epsilon|X,\beta) = f_{\varphi}(\epsilon^{-1}(\varphi)\,|X,\beta) \left|\frac{d\varphi}{d\epsilon}\right| = exp(\epsilon - exp(\epsilon))$$

so that ϵ has an extreme value distribution. This transformation result is identical to that derived for the case of non-time varying covariates considered by HH. This specification of the proportional hazards model and that given in (1.3) differ in that the linear regression form for the impact of the covariates does not hold when the latter are allowed to vary over time. To see this, return to the general specification given in (1.2) where the unconditional density function depends upon time-varying covariates $X(t)$. Then, using the definitions given in (1.4) and (1.6), the model can be written as

$$(1.8) \qquad f(t\,|X,\beta) \;=\; Pr(log\ (\int_0^t \lambda_o(s)\ exp(X(s)\beta)\ ds) \;=\; \epsilon)$$

where ϵ again follows the extreme value distribution. Since both the hazard and the covariates vary over time, the strong separability observed in (1.3) does not hold. Note that if $X(t)$ is constant then (1.8) does simplify to the time-invariant form.

Since all of the terms in the integral of (1.8) are positive, the integral is monotonically increasing in t so that the probability that the failure time is less than t is given by

$$(1.9) \qquad F(t\,|X,\beta) \;=\; Pr(log\ (\int_0^t \lambda_o(s)\ exp(X(s)\beta)\ ds) \;+\; X\beta \;\geq\; \epsilon)$$

The probability that the failure time is in the interval $(t-1,t]$ is simply $F(t\,|X,\beta) - F(t-1\,|X,\beta)$.

In order to estimate this model semi-parametrically, it is necessary to take advantage of the discrete nature of the data by placing conditions on the evolution of the covariates over time. Specifically, I will assume that the time-varying covariates $X(t)$ are constant over intervals defined in discrete time. More formally, if there are T potential failure times, then $X(\text{j}-\Delta) = X(\text{j})$; $\Delta \in (0,1)$, j $= 1,\ldots,T$.

The assumption that variables are constant over a discrete interval is quite reasonable for most applications. Economic data that are collected over time are invariably sampled at discrete intervals. While some financial data are collected as often as daily or even hourly, standard survey data are generated more infrequently, generally yearly. Even for the relatively rare hourly data, the observation interval is discrete in the sense that the data do not include observations on the *continuous* evolution of the variables. The assumption that covariates are constant over the observation interval is a natural one given the inherent discreteness of sampling and the lack of *a priori* knowledge about the path of the data between discrete observations over time.[2]

Under these assumptions, the specification of the likelihood of an observation on a failure time is straightforward. For notational convenience, define the following terms:

(1.10)
$$\tau(j) \equiv \int_{j-1}^{j} \lambda_o(s)\, ds,$$
$$Z(j) \equiv exp(X(j)\beta),$$
$$\Psi_r \equiv log\ (\sum_{i=1}^{r} Z(i)\tau(i)),$$

r = 1,2,...,T, where the constancy of $X(t)$ over an interval implies

---

[2]The assumption of time-varying variables that are constant over an interval is not actually requried. So long as one is willing to make an assumption about the evolution of the variables over the interval, then the identification results below will follow. Alternative specifications of this process may, however, complicate the likelihood function considerably.

that $Z(s-\Delta) = Z(s)$, $s = 1,\ldots,T$; $\Delta \in (0,1)$. For expositional purposes, failure times are assumed to occur in the interval prior to the recorded failure time so that a failure time of t represents failure in the half-open interval $(t-1,t]$. Then denoting $\tau' = (\tau(1),\cdots,\tau(T))$, the likelihood of observing a failure in the discrete interval $(t-1,t]$ is given by

$$(1.11) \qquad L_t(\beta,\tau|X) = \int_{\Psi_{t-1}}^{\Psi_t} f(\epsilon)\, d\epsilon = F(\Psi_t) - F(\Psi_{t-1}),$$

while the likelihood of observing a censored failure at t is

$$(1.12) \qquad L_t^c(\beta,\tau|X) = \int_{\Psi_t}^{\infty} f(\epsilon)\, d\epsilon = 1 - F(\Psi_t).$$

Inspection of the likelihood functions (1.11) and (1.12) reveals a close correspondence between the Ψ-functions and the ℓ-functions found in HH. In this more general specification, the constant terms in the ordered logit are not simply functions of the baseline hazard, but rather are functions of weighted sums of the discrete segments of the baseline hazard, with the weights corresponding to the effects of the time-varying covariates. If there are time-constant covariates, they enter the function as a standard separate linear term rather than as weights.

Consider an independent sample of the form (y_i,X_i,δ_i) i=1,...,N, where y_i are indicators on failure for individual i for each time period, X_i are individual and time-specific vectors of covariates, and δ_i are censoring indicators for each time period. Then the log likelihood of the sample is given by:

$$(1.13) \qquad \log L_N(\beta,\tau) = \sum_{i=1}^{N} \sum_{t=1}^{T} \Big[y_{it}(1-\delta_{it}) \log(F(\tau_t^i) - F(\tau_{t-1}^i)) + \delta_{it} \log(1 - F(\tau_t^i)) \Big] .$$

The addition of unobserved individual heterogeneity to the model is straightforward and requires only one additional order of integration. If the heterogeneity is of standard multiplicative form ($\lambda(t|X,\beta) = \theta\ \lambda_o(t)\ exp(X(t)\beta)$) then all that is required to add heterogeneity to the model is to replace the distribution function $F(\epsilon)$ with $E_\theta\ F_{\epsilon+\theta}(\epsilon+\theta|\theta)$. It is easy to derive a closed-form expression for this likelihood in the case where θ has a gamma distribution (Appendix 2).

1.2 Identification, Consistency and Asymptotic Normality

Under suitable regularity conditions, the maximum likelihood estimator of the likelihood function (1.13) is consistent and asymptotically normal. Given the close relationship between the ordered logit form of this likelihood and the likelihoods associated with the general multinomial response model, the following discussion follows closely the existing discrete choice literature. All results assume a finite number of periods T.

The proof of consistency and asymptotic normality involves verifying the conditions discussed in McFadden [1984] and is completed in three stages. First I show that the identification condition (4) of McFadden is implied by showing that the time interval probabilities are uniquely determined, period by period. Second, I demonstrate that under rather weak regularity conditions on the data, τ and β uniquely determine the *CDF* evaluation points, and hence the interval probabilities for all observations. For the case where there is individual heterogeneity, slightly stronger conditions are required. Finally, I verify that the other necessary assumptions for consistency and asymptotic normality are satisfied.

Consider a sample log likelihood function of the form given in (1.13). I make the following initial assumptions:

A1.1 - Error Distribution -

The distribution function F is continuous and twice differentiable with density function $f > 0$ almost everywhere.

A1.2 - Covariates -

The domain of the explanatory variables is a measurable subset of $\mathbb{R}^k$ with probability $P(x)$, where k is an integer such that $k < \infty$.

A1.3 - Parameters -

The true parameter vector $\theta_o = (\beta_o, \tau_o)$ is an interior point of the parameter space, the compact set $B \times L$, where $B \subset \mathbb{R}^k$, $L \subset \mathbb{R}^T$.

These relatively innocuous conditions correspond to regularity conditions (1), (2), (3) and (5) of McFadden [1984]. In general, the present assumptions are stronger than their counterparts.

To insure identification, consistency and asymptotic normality, additional conditions on the data are required:

A1.2 Covariates

Add to the existing assumptions,

(i) *boundedness* - $\exists\, M < \infty$, *s.t.* $E|X|^3 < M$

(ii) *non-degeneracy* - for $\forall\, c \in \mathbb{R}$, $\beta \in \mathbb{R}^k$ $(\beta \neq 0)$, $\exists$ t, *s.t.*

$$Pr_X(X(t)\beta = c) < 1.$$

A1.4 Information

The information matrix evaluated at the true parameter vector, defined as,

$$J(\theta_o) \equiv \int dP(x) \sum_{t=1}^{T} Pr^t(x;\theta_o) \left[\frac{\partial log\ Pr^t(x;\theta_o)}{\partial \theta} \right] \left[\frac{\partial log\ Pr^t(x;\theta_o)}{\partial \theta} \right]'$$

is non-singular.

Assumption A.1.2.(i) implies that McFadden condition (6) holds. Assumption A1.2.(ii) is essentially a linear independence condition on the X. It is analagous to the non-singularity of the design matrix in least squares estimation and is a natural generalization of the HH condition (A1.2) to the case of time-varying covariates. Note, however, that this condition requires only that the covariates associated with at least one period have a non-degenerate distribution, rather than placing restrictions upon the behavior in every period. The information matrix assumption is self-explanatory.

Note that the likelihood function given in (1.13) may be written as a special case of the general multinomial response model:

$$L(\beta, \tau) = \sum_{i=1}^{N} \sum_{m=1}^{M} y_{im} \, log \; Pr^{m}(X_{it}, \beta) \tag{1.14}$$

where the y_{im} are indicator functions on individuals i and categories m, and the Pr^{m} are the probabilities defined over the m categories. In this particular application, the categories are defined by the failure time intervals.[3]

The correspondence between the ordered logit model and the

[3]In the preceeding discussion, censoring is ignored. The addition of independently censored observations does not change the identification results, but requires somewhat more complex notation to normalize the probabilities.

general case considered by McFadden implies that for this specification, the identification condition is given by (4) of McFadden-

(4 - McFadden) For $\forall\ \epsilon > 0,\ \exists\ \delta < 0$ such that $|\theta_o - \theta| \geq \epsilon$ implies:

$$\int dP(x) \sum_{t=1}^{T} Pr^t(x;\theta_o)\ log\left[\frac{Pr^t(x;\theta)}{Pr^t(x;\theta_o)}\right] < \delta.$$

Intuitively, the condition requires that there be no parameter vector that achieves as high a limiting value of the log likelihood as the true value.

There is a restatement of this condition which makes the demonstration of identication somewhat easier. Using the fundamental information inequality, Rao (1.e.6), and taking advantage of the fact that the probabilities are exhaustive over the t categories so that $\Sigma\ Pr^t(X;\theta_o) = \Sigma\ Pr^t(X;\theta)\ \forall\ \theta$, condition (4-McFadden) implies that:

$$(1.15) \qquad \int dP(x) \sum_{t=1}^{T} Pr^t(x;\theta_o)\ log\left[\frac{Pr^t(x;\theta)}{Pr^t(x;\theta_o)}\right]$$

$$= \qquad - \int dP(x) \sum_{t=1}^{T} \frac{(Pr^t(x;\theta_o) - Pr^t(x;\theta))^2}{Pr^t(x;\theta_o)\quad 2\ g(x)^2}$$

which is ≤ 0.

Since the terms in the numerator and denominator are non-negative, given continuity on the Pr^t in θ for $\forall X$, equality holds if each of the category probabilities are identical under θ and θ_o for X with positive probability.[4] This implies that demonstrating the identification condition for the ordered logit model requires showing that conditional upon X, the *CDF* evaluation points are uniquely determined by choice of β and τ. By standard arguments, demonstrating that no other parameter vector yields as high a value of the continuous in θ limiting log likelihood implies A1.4, the non-singularity of $J(\theta)$.

The proof of the uniqueness of the intervals is straightforward. To prove uniqueness, I will make use of the following induction lemma:

Lemma 1.1 - If the parameters for the first K-1 intervals are identified, then the additional parameter for the K-th interval is also identified.

Proof: Suppose that the first K-1 intervals are identified. Then in considering the K-th interval it suffices to restrict the relevant comparison of parameters to the subspace of the parameter space for

---

[4]Continuity of the Pr^t in θ follows from the continuity assumption on the underlying f and F functions and the definition of the Ψ-functions.

which the $\omega_{K-1} \equiv (\beta, \tau(1), \tau(2), \cdots, \tau(K-1))$ are equal.[5] Consider two distinct sets of parameters $\Gamma_1 = (\bar{\omega}_{K-1}, \tau_1(K))$ and $\Gamma_2 = (\bar{\omega}_{K-1}, \tau_2(K))$, for any valid choice of $\bar{\omega}_{K-1} = (\bar{\beta}, \bar{\tau}(1), \bar{\tau}(2), \cdots, \bar{\tau}(K-1))$. Changing notation slightly, recall that

$$\Psi_K(\theta) = \Psi_K(\omega_{K-1}, \tau(K)) = \log\left(\sum_{s=1}^{K} \exp(X(s)\beta)\tau(s)\right).$$

Then the probability of Γ_1 and Γ_2 yielding the same Ψ_K is

$$\begin{aligned} &Pr_X\Big[\log[\Psi_{K-1}(\bar{\omega}_{K-2}, \bar{\tau}(K)) + \exp(X(s)\bar{\beta})\tau_1(K)] \\ &\quad = \log[\Psi_{K-1}(\bar{\omega}_{K-2}, \bar{\tau}(K)) + \exp(X(s)\bar{\beta})\tau_2(K)] \Big] \\ &= Pr_X(\tau_1(K) = \tau_2(K)), \end{aligned}$$

which equals 0 since $\tau_1(K) \neq \tau_2(K)$ by assumption. Thus, provided that the K-1 preceeding intervals are identified, the K-th interval is identified.

The identification lemma follows directly from this result.

---

[5]Consideration of the unrestricted space is not necessary since the identification of the parameters for the first K-1 periods implies that the parameters for the K-1 periods are identical.

Lemma 1.2 - Given A1.1-A1.3, the interval probabilities are uniquely determined by choice of β and τ.

Proof: For this model, the uniqueness of the interval probabilities is equivalent to the uniqueness of the evaluation points. Without loss of generality, assume that the first K-1 periods have degenerate X's and that uniqueness of the β and τ-vectors is not guaranteed. Define the terms $c_t(\beta,\tau) \equiv exp(X(t)\beta)\,\tau(t)$, t = 1,...,T. From the definitions above, it follows that $\Psi_t(\beta,\tau) = log(\sum_{s=1}^{t} c_s(\beta,\tau))$.

Since the first K-1 periods are degenerate, it is possible to have two sets of parameters which yield the same evaluation points for each period. Thus, for $\omega^{o}_{K-1} \equiv (\beta_o,\tau_o(1),\cdots,\tau_o(K-1))$ and $\omega^{1}_{K-1} \equiv (\beta_1,\tau_1(1),\cdots,\tau_1(K-1))$ each of the evaluation points is identical so that $c_t(\beta_o,\tau_o) = c_t(\beta_1,\tau_1)$, t = 1,...,K-1. Then, defining $S_{K-1}(\beta,\tau) \equiv \sum_{t=1}^{K-1} c_t(\beta,\tau)$, it follows that $S_{K-1}(\beta_o,\tau_o) = S_{K-1}(\beta_1,\tau_1)$.

Now consider the K-th period. Rewrite $\Psi_K(\beta,\tau) = log(S_{K-1}(\beta,\tau) + c_K(\beta,\tau))$. Since S_{K-1} must be identical across any possible parameter comparisons, it suffices to demonstrate the identification of c_K. Consider two sets of parameters $\Gamma_1 = (\beta_1,\tau_1)$, and $\Gamma_2 = (\beta_2,\tau_2)$ which are drawn from a subspace of the parameter space that yields $c_t(\beta_1,\tau_1) = c_t(\beta_2,\tau_2)$, t = 1,...,K-1. Then the probability that the two sets of parameters yield the same evaluation point is given by:

$$Pr_X(exp(X(K)\beta_1)\tau_1(K) = exp(X(K)\beta_2)\tau_2(K))$$
$$= \quad Pr_X(X(K)\beta^* = \kappa)$$

where κ is a constant equal to $-log(\tau_1(K)/\tau_2(K))$. By A1.2, this probability is strictly less than 1 for all choice of β_1 and β_2 so that the β and $\tau(K)$ are identified.

Since β is identified, it follows immediately that $\tau(1)$ is identified; the conditions $c_1(\beta_1,\tau_1) = c_1(\beta_2,\tau_2)$ and $\beta_1 = \beta_2$ imply trivially that $\tau_1(1) = \tau_2(1)$. Thus, the parameters for the first interval are identified. Application of the Lemma 1.1 implies that the parameters for every period are identified.

Under suitable regularity conditions standard theory can be applied to derive the asymptotic distribution of the maximum likelihood estimator of θ. The asymptotic results assume that the number of intervals T is fixed while the sample size $N \to \infty$. The results are summarized in the following theorem:

Theorem 1.1 - Given A1.1-A1.3, then almost surely a unique maximum likelihood estimator satisfying $\hat{\theta}_N \equiv (\hat{\beta}_N,\hat{\tau}_N)$, $\partial L_N(\theta)/\partial\theta = 0$ eventually exists and $\hat{\theta}_N \to \theta_o$. Furthermore, $\sqrt{N}(\hat{\theta} - \theta_o)$ $\overset{a}{\sim} N(0,J(\theta_o)^{-1})$.

Proof: A1.1-A1.3 and the Lemmas 1.1 and 1.2 imply that conditions 1-6 of McFadden are satisfied so that the McFadden Theorem 2 demonstrating consistency applies. A1.4 is equivalent to regularity condition 7 of McFadden and follows given A1.1-A1.3. Asymptotic normality follows by Theorem 3 of McFadden. It is easy to verify that the ordered extreme value model satisfies the assumptions in A1.1. I assume that the data and parameters satisfy A1.2, A1.3.

Consider now the case where there is unobserved individual heterogeneity. The proofs follow closely HH with allowance made for the different assumptions and more complicated notation needed for the extension to time-varying covariates. Assume that the heterogeneity is of the standard multiplicative form and that $E_\theta\, F_{\epsilon+\theta}(\epsilon+\theta|\theta)$ can be rewritten as $F_\mu(\mu|\gamma)$, where γ is a finite dimensional parameter vector.

To demonstrate the identification condition, it is necessary to show that there do not exist different combinations of (β,τ,γ) parameters that will yield the same *CDF* values for all possible observations on the covariates. Note that in the absence of heterogeneity, it was necessary only to show that the evaluation points are the same. Since the shape of the distribution function depends upon the parameter vector γ, it is necessary here to consider the actual probability values generated by the covariates. This requires substantive modifications in the assumptions given in

A1.1-A1.3:

A1.1' - Error Distribution -

(i) The distribution function for the error term μ, $F(\mu|\gamma)$, depends upon an $m \times 1$ finite dimensional parameter vector γ, is twice differentiable in μ and γ, and has density function $f(\mu|\gamma) > 0$ almost everywhere, $\forall\ \gamma \in \mathbb{R}^m$.

(ii) The set $\{\mu | f(\mu|\gamma) = a\ f(\frac{b + \mu}{c}|\gamma_o)\}$ has Lebesgue measure zero $\forall\ \gamma \in \mathbb{R}^m$ and $a,b,c \in \mathbb{R}$.

A1.2' - Covariates-

In addition to A1.2.(i) and A1.2.(ii) the following assumption is made about the covariates:

(iii) For some $t \in \{1,...,T\}$ satisfying A1.2.(ii), $\exists\ h \in \{1,...,k\}$ *s.t.* $\beta_h \neq 0$ and, conditional on $\tilde{X}_h(t)$, the vector of covariates with the h-th element of the t-th period deleted, $X_h(t)$ has positive density in an open neighborhood in $\mathbb{R}$ *a.s.* P_X.

A1.3' - Parameters -

The true parameter vector $\theta_o=(\beta_o,\tau_o,\gamma_o)$ is an interior point of the parameter space, the compact set $B \times L \times M$ where $B \subset \mathbb{R}^k$, $L \subset \mathbb{R}^T$, $M \subset \mathbb{R}^m$.

The assumption A1.1.(i)' places conditions upon the conditional distribution function analagous to those in A1.1.(i) which insure that it is well behaved for all values of γ. The additional assumption on the density function, A1.1.(ii)' is used in conjunction with the additional condition on the covariates A1.2.(iii)' to guarantee that there is enough variability and continuity in the data across observations and in the density function over different values of γ so that the evaluation points can be separated from the γ parameter.[6] Note that the t in A1.2.(iii)' refers to the same period as the t in A1.2.(ii).

Under these assumptions, the following lemma demonstrates uniqueness of the interval probabilities:

Lemma 1.3 - Given A1.1-A1.3, the interval probabilities are uniquely determined by choice of β, τ and γ.

Proof: I will restrict the discussion to the case where $\gamma \neq \gamma_o$ since if $\gamma = \gamma_o$ the proof from above applies. For $\gamma \neq \gamma_o$ is sufficient to show that there does not exist $\theta \neq \theta_o$, such that for $\forall\ t$, $Pr_X\ [F(\varphi_t(\theta))) = F(\varphi_t(\theta_o))] = 1$. In particular, this

---

[6]The assumption A1.1(ii)' rules out cases like a mixture of a normal error terms with normal heterogeneity. For obvious reasons, it is impossible to separate out the heterogeneity parameter from the underlying β and τ in a normal mixture model.

relationship must hold for a period s satisfying A1.2.(iii)'. Suppose that such a θ exists. Define $\tilde{X}(s)$ and $\tilde{\beta}$ which are the appropriate parts of the $X(s)$ and β vectors with the h-th element deleted, and the evaluation point

$$R_s(\theta, X, \Delta) = log\left[exp(\Psi_{t-1}(\theta)) + exp\{\tilde{X}(s)\tilde{\beta} + (X_h(s)+\Delta)\beta_h\}\ \tau(s)\right].$$

Note that $R_s(\theta, X, 0) = \Psi_s(\theta)$. Since $X(s)$ is assumed not to be degenerate, $F_s(R_s(\theta, X, 0)) = F_s(R_s(\theta_o, X, 0))\ \forall\ X$, and F is assumed to be continuous and differentiable, it follows that

$$\lim_{\Delta \to 0} \frac{F(R_t(\theta, X, \Delta)\,|\gamma) \quad - \quad F(\Psi_t(\theta)\,|\gamma)}{\Delta}$$

$$= \quad \lim_{\Delta \to 0} \frac{F(R_t(\theta_o, X, \Delta)\,|\gamma_o) \quad - \quad F(\Psi_t(\theta_o)\,|\gamma_o)}{\Delta}$$

for $\forall\ X_h(s)$ in the open neighborhood of $\mathbb{R}$ discussed in A1.2.(iii)'. Then

$$f(\Psi_s(\theta)\,|\gamma) = a\ f(\Psi_s(\theta)+b\,|\gamma_o)$$

where $$a = \frac{\beta_{oh}\ exp(X(s)\beta_o)\ \ \tau_o(s)}{\beta_h\ exp(X(s)\beta)\ \ \tau(s)}\ \ \frac{exp(\Psi_s(\theta))}{exp(\Psi_s(\theta_o))}$$

and $$b = \Psi_s(\theta_o) - \Psi_s(\theta).$$

By A1.2.(iii)', the set of $X(s)$'s which satisfy this relationship has positive Lebesgue measure so that the set of Ψ-evaluation points satisfying this relationship has positive measure, violating A1.1.(ii)'.

The consistency and asymptotic normality results follow immediately:

Theorem 1.2 - Given A1.1'-A1.3', then almost surely a unique maximum likelihood estimator satisfying $\hat{\theta}_N \equiv (\hat{\beta}_N, \hat{\tau}_N, \hat{\gamma}_N)$, $\partial L_N(\hat{\theta})/\partial\theta = 0$ eventually exists and $\hat{\theta}_N \to \theta_o$. Furthermore, $\sqrt{N}(\hat{\theta} - \theta_o) \overset{a}{\sim} N(0, J(\theta_o)^{-1})$.

Proof: A1.1'-A1.3' and the Lemmas 1.1 and 1.2 imply that the consistency conditions are satisfied. Since A1.4 holds under A1.1'-A1.3', the model satisfies the asymptotic normality conditions.

2. Competing Risks Models

2.1 Specification

The competing risks model can be cast in terms of the familiar latent variables specification. Suppose that there are random variables for J different potential failure times T_j, $j = 1,...J$, where only the first realized failure time is observed so that $T = min_j \{T_1,...,T_J\}$. I assume the cause-specific hazard rates at time period t for cause j to be of the proportional hazards form

$$\text{(2.1)} \qquad \lambda_j(t|X_j(t),\beta_j) = \lambda_{jo}(t)\ exp(X_j(t)\beta_j).$$

Ruling out simultaneous failures, the overall hazard is simply $\Sigma_j\ \lambda_j$.

Kalbfleisch and Prentice [1984] point out that the standard likelihood function factors into separate components for each risk where failures of an alternative type are treated as censored observations. This implies that the transformation to the regression form of the likelihood function given above is valid for competing risks. First define the cause-specific analogues to the terms in (1.9):

$$\text{(2.2)} \qquad \tau_j(r) \equiv \int_{r-1}^{r} \lambda_{jo}(s)\ ds,$$

$$Z_j(r) \equiv exp(X_j(r)\beta_j),$$

$$\gamma_r^j \equiv log\ (\sum_{i=1}^{r} Z_j(i)\tau_j(i)),$$

for $j = 1,...,J$. Then the unconditional density for a failure of type j at time t can be written in the equivalent form $\Psi_t^j = \epsilon_j$ where the error terms ϵ_j are independent extreme value errors.

Specializing to the case of bivariate independent competing risks where data are again observed at discrete intervals, the specification of a likelihood function is similar to the case for non-time varying covariates. First, consider the simple case of an observation censored at time period t:

$$(2.3) \qquad L_t^c(\beta,\tau\,|\,t,X) = \int_{\Psi_t^1}^{\infty} \int_{\Psi_t^2}^{\infty} f(\epsilon_1,\epsilon_2)\; d\epsilon_2\; d\epsilon_1.$$

With discrete data, the likelihood of an observed failure of a given type is somewhat more involved because the integration must insure that the latent failure time is greater than the observed failure time for every point in the discrete interval. If, without loss of generality, a failure of type 1 is observed in the interval $(t-1,t]$, the likelihood is given by

$$(2.4) \qquad L_t^1(\beta,\tau\,|\,X) = \int_{\Psi_{t-1}^1}^{\Psi_t^1} \int_{g(\epsilon_1)}^{\infty} f(\epsilon_1,\epsilon_2)\; d\epsilon_2\; d\epsilon_1.$$

where the $g(\cdot)$ function insures that the required relationship between the failure times holds. Then for a failure of type 1, the functional form of $g(\cdot)$ is given by:

$$(2.5) \qquad g(\mu) = \psi^2_{t-1} + (\mu - \psi^1_{t-1}) \frac{\psi^2_t - \psi^2_{t-1}}{\psi^1_t - \psi^1_{t-1}} .$$

The derivation of $g(\cdot)$ involves the assumption that the evaluation points change at a constant rate over the discrete intervals.[7] Consider a failure time t^* which lies in the interval $(t-1, t]$. Then given the assumption of linearity in the change in evaluation points over the interval,

$$(2.6) \qquad \epsilon^*_j \equiv \psi^j_{t^*} = \psi^j_{t-1} + (\psi^j_t - \psi^j_{t-1}) \ (t^* - (t-1))$$

j=1,2. Solving for $(t^*-(t-1))$ for both risks and setting the terms equal to each other yields

$$(2.7) \qquad \frac{\epsilon^*_1 - \psi^1_{t-1}}{\psi^1_t - \psi^1_{t-1}} = \frac{\epsilon^*_2 - \psi^2_{t-1}}{\psi^2_t - \psi^2_{t-1}}$$

and solving for ϵ^*_2 in terms of ϵ^*_1 yields the desired result.

Suppose that an independent sample on data of the form

---

[7] This assumption is analagous to the one made by Han and Hausman. The functional form of the g-function implied by the linearity is somewhat more complicated because the the non-separability of the covariates from the hazard. Note, however, that in the case of constant covarariates, the likelihood reduces to the HH case.

$(y_i^1, y_i^2, X_i, \delta_i)$ i=1,...,N, are observed where y_i^k are indicators on failure of type k for individual i at each period, X_i are individual and time-specific vectors of covariates, and δ_i are censoring indicators for each time period. Then the log likelihood of the sample is:

$$(2.8) \qquad \log L_N(\beta, \tau) = \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\sum_{k=1}^{2} y_{it}^k (1-\delta_{it}) \log L_{it}^k \right] + \delta_{it} \log L_{it}^c .$$

As in the single risk case, it is possible to replace the density function f in (2.3) and (2.4) with a density function that depends upon a finite dimensional parameter $f(\cdot|\gamma)$. The additional parameter can be thought of as arising from either unobserved heterogeneity, correlation between the underlying risks or some combination of the two. The assumption adopted by HH is to assume that $f(\cdot|\gamma)$ is a bivariate normal where γ is the correlation coefficient ρ.

2.2 Identification, Consistency and Asymptotic Normality

As with the single risk case, the likelihood (2.8) is a special case of the general multinomial response model with the categories defined by the failure type and time interval combinations. As a result, demonstrating identification involves showing that the probabilities associated with each failure time and type combination are unique with respect to choice of $\beta_1, \beta_2, \tau_1, \tau_2$ and ρ.

The proof will be undertaken in three steps. In the first two, I demonstrate that the identification condition holds. In particular, I show that for a given set of probabilities associated with each failure time and type, the evaluation points in the bivariate distribution function are uniquely determined. I then show that given realizations on the covariates the set of evaluation points is unique with respect to choice of the parameter vector. Finally I verify the additional conditions required for consistency and asymptotic normality. I make the following assumptions:

A2.1 - Error Distribution -

The distribution function F is continuous and twice differentiable with density function $f > 0$ almost everywhere.

A2.2 - Covariates -

The domain of the explanatory variables is a measurable subset of $\mathbb{R}^k$, with probability $P(x)$ where $k=k_1+k_2$, and k_1, k_2 are integers such that $k_1, k_2 < \infty$. The following conditions are also satisfied:

(i) *boundedness* - $\exists\, M < \infty$, *s.t.*

$$E|X|^3 < M,$$

(ii) *non-degeneracy* - for $\forall\ c \in \mathbb{R}$, $\beta_1 \in \mathbb{R}^{k_1}$ $(\beta_1 \neq 0)$, $\beta_2 \in \mathbb{R}^{k_2}$ $(\beta_2 \neq 0)$, $\exists\ t$ and $i \in \{1,2\}$, *s.t.*

$$Pr_X\ (X_i(t)\beta_i = c) < 1,$$

where $X_i(t)$ is the vector of covariates for risk i at time period t.

A2.3 - Parameters -

The true parameter vector $\theta_o=(\beta_{1o},\beta_{2o}\tau_{1o},\tau_{2o})$ is an interior point of the parameter space, the compact set $B_1 \times B_2 \times L_1 \times L_2$, where $B_1 \subset \mathbb{R}^{k_1}$, $B_2 \subset \mathbb{R}^{k_2}$, $L_1, L_2 \subset \mathbb{R}^T$.

The first step is to show that a given set of failure type/time probabilities yields unique interval evaluation points. Recall that the probabilities are given by functions of the form (2.4). Consider two adjacent evaluation points, (γ_t^1, γ_t^2) and $(\gamma_{t-1}^1, \gamma_{t-1}^2)$. I will first show for a given period, that if two sets of non-identical evaluation points yield the same interval probabilities, then the evaluation points for both periods differ. This result can be used to show that the evaluation points are uniquely determined by the choice-specific probabilities. More formally,

Lemma 2.1 - Suppose that there exist two sets of evaluation points $\{(\gamma_{t-1}^1, \gamma_{t-1}^2), (\gamma_t^1, \gamma_t^2)\}$ and $\{(\tilde{\gamma}_{t-1}^1, \tilde{\gamma}_{t-1}^2), (\tilde{\gamma}_t^1, \tilde{\gamma}_t^2)\}$ that yield the same choice-specific probabilities defined over $(t-1,t]$. Then $(\psi_t^1,\psi_t^2) \neq (\tilde{\psi}_t^1,\tilde{\psi}_t^2)$ implies that $(\psi_{t-1}^1,\psi_{t-1}^2) \neq (\tilde{\psi}_{t-1}^1,\tilde{\psi}_{t-1}^2)$.

Proof: Suppose that the two sets of evaluation points discussed above exist and that $(\psi_{t-1}^1,\psi_{t-1}^2) = (\tilde{\psi}_{t-1}^1,\tilde{\psi}_{t-1}^2)$. Let $\psi_{t-1}^1 = \tilde{\psi}_{t-1}^1 = c_1$ and $\psi_{t-1}^2 = \tilde{\psi}_{t-1}^2 = c_2$ for some constants $c_1, c_2 \in \mathbb{R}$. There are no restrictions on the underlying parameters (β,τ) and $(\tilde{\beta},\tilde{\tau})$. Without loss of generality, let $\tilde{\psi}_t^1 > \psi_t^1$. Since the interval probabilities are assumed to be the same,

$$(2.9) \qquad \int_{c_1}^{\psi_t^1} \int_{g_1(\epsilon_1)}^{\infty} f(\epsilon_1,\epsilon_2)\, d\epsilon_2\, d\epsilon_1 = \int_{c_1}^{\tilde{\psi}_t^1} \int_{h_1(\epsilon_1)}^{\infty} f(\epsilon_1,\epsilon_2)\, d\epsilon_2\, d\epsilon_1$$

where

$$g_1(\mu) = c_2 + (\mu - c_1)\,\frac{\psi_t^2 - c_2}{\psi_t^1 - c_1}$$

$$h_1(\mu) = c_2 + (\mu - c_1)\,\frac{\tilde{\psi}_t^2 - c_2}{\tilde{\psi}_t^1 - c_1}$$

The equality in (2.9) and assumption that f has positive density almost everywhere implies that $h_1(\cdot) > g_1(\cdot)$. This in turn requires that

$$\text{(2.10)} \qquad \frac{\tilde{\psi}_t^2 - c_2}{\tilde{\psi}_t^1 - c_1} > \frac{\psi_t^2 - c_2}{\psi_t^1 - c_1}$$

and

$$\tilde{\psi}_t^2 > \psi_t^2 + (\tilde{\psi}_t^1 - \psi_t^1)\, \frac{\psi_t^2 - c_2}{\psi_t^1 - c_1} \,,$$

The last inequality guarantees that $\tilde{\psi}_t^2 > \psi_t^2$. The type 2 interval probabilities are of the form

$$\text{(2.11)} \qquad \int_{c_2}^{\tilde{\psi}_t^2} \int_{h_2(\epsilon_2)}^{\infty} f(\epsilon_1, \epsilon_2)\, d\epsilon_1\, d\epsilon_2,$$

where h_2 is of the same form as h_1 but with the indices reversed. With some manipulation of the integrals, (2.11) can be rewritten as

$$\text{(2.12)} \qquad \int_{c_2}^{\psi_t^2} \int_{g_2(\epsilon_2)}^{\infty} f(\epsilon_1, \epsilon_2)\, d\epsilon_1\, d\epsilon_2$$

$$+ \int_{\psi_t^2}^{\tilde{\psi}_t^2} \int_{h_2(\epsilon_2)}^{\infty} f(\epsilon_1, \epsilon_2)\, d\epsilon_1\, d\epsilon_2$$

$$+ \int_{c_2}^{\tilde{\psi}_t^2} \int_{g_2(\epsilon_2)}^{h_2(\epsilon_2)} f(\epsilon_1, \epsilon_2)\, d\epsilon_1\, d\epsilon_2$$

where g_2 is analagous to g_1. Since $\tilde{\gamma}_t^2 > \gamma_t^2$ the second term in (2.12) is strictly greater than 0 so that the equality of (2.11) and (2.12) holds if and only if $h_2 < g_2$. But this inequality contradicts (2.10). Thus, if $(\gamma_t^1, \gamma_t^2) \neq (\tilde{\gamma}_t^1, \tilde{\gamma}_t^2)$, a necessary condition for probability interval equality over $(t-1, t]$ is $(\gamma_{t-1}^1, \gamma_{t-1}^2) \neq (\tilde{\gamma}_{t-1}^1, \tilde{\gamma}_{t-1}^2)$.

This result implies that the entire sequence of evaluation points is unique:

Lemma 2.2 - If two sets of evaluation points $\{(\gamma_1^1, \gamma_1^2), \ldots, (\gamma_T^2, \gamma_T^2)\}$ and $\{(\tilde{\gamma}_1^1, \tilde{\gamma}_1^2), \ldots, (\tilde{\gamma}_T^1, \tilde{\gamma}_T^2)\}$ yield the same set of choice-specific probabilities, then $\gamma_t^j = \tilde{\gamma}_t^j$, $t = 1, \ldots, T$, $j = 1, 2$.

Proof: Without loss of generality, suppose that there is an evaluation point $(\gamma_t^1, \gamma_t^2) \neq (\tilde{\gamma}_t^1, \tilde{\gamma}_t^2)$ where the interval probabilities are the same. Then working backward by repeated application of Lemma 2.1, the first interval evaluation points differ, $(\gamma_1^1, \gamma_1^2) \neq (\tilde{\gamma}_1^1, \tilde{\gamma}_1^2)$. But this result yields a contradiction since $(\gamma_0^1, \gamma_0^2) = (\tilde{\gamma}_0^1, \tilde{\gamma}_0^2) = (-\infty, -\infty)$.

The final lemma demonstrates that under regularity conditions on the distribution of the covariates, a set of evaluation points uniquely determines the β, τ.

Lemma 2.3 - Given A2.1-A2.3, the choice-specific interval probabilities are uniquely determined by choice of $\beta=(\beta_1,\beta_2)$ and $\tau=(\tau_1,\tau_2)$.

Proof: Lemmas 2.1-2.2 imply that showing uniqueness of the evaluation points is sufficient to demonstrate uniqueness of the probabilities. Note that the choice-specific evaluation points for each risk are of the same form as the single risk evaluation points in section 1. A2.1-A2.3 satisfy the conditions of Lemma 1.3 for each risk, therefore application of Lemma 1.3 to either risk provides the uniqueness result.

These results imply that under rather weak assumptions the standard consistency and asymptotic normality results hold. The results are described in the following theorem:

Theorem 2.1 - Given A2.1-A2.3, then almost surely a unique maximum likelihood estimator satisfying $\hat{\theta}_N \equiv (\hat{\beta}^1_N,\hat{\beta}^2_N,\hat{\tau}^1_N,\hat{\tau}^2_N)$, $\partial L_N(\hat{\theta})/\partial\theta = 0$ eventually exists and $\hat{\theta}_N \to \theta_o$. Furthermore, $\sqrt{N}(\hat{\theta} - \theta_o) \overset{a}{\sim} N(0,J(\theta_o)^{-1})$.

Proof: A2.1-A2.3 and the lemmas imply that the consistency conditions of McFadden Theorem 2 are satisfied. The assumptions imply that A1.4 is satisfied so that the model satisfies the asymptotic normality conditions of McFadden Theorem 3.

Finally, I consider the case where the distribution function depends upon a finite-dimensional parameter γ. As with the single risk case, the identification of this model requires stronger conditions upon the covariates and the distribution function of the error. I make the following modifications to the assumptions given in A2.1-A2.3:

A2.1' - Error Distribution -

(i) The distribution function for the error term $F(\cdot,\cdot|\gamma)$ depends upon an $m \times 1$ finite dimensional parameter vector γ, is continuous and twice differentiable in its arguments and in γ, and has density function $f > 0$ almost everywhere $\forall\ \gamma \in \mathbb{R}^m$.

(ii) One of the following applies:

(a) If the covariates differ across risks so that $X_1 \neq X_2$ then

The set $\{(\mu_1,\mu_2)\,|f(\mu_1,\mu_2|\gamma) = a\ f(\dfrac{\mu_1+b_1}{c_1},\dfrac{\mu_2+b_2}{c_2}\ |\gamma_o)\}$ has $\mathbb{R}^2$-Lebesgue measure zero for $\gamma \neq \gamma_o$, $\forall$ $a,b_1,b_2,c_1,c_2 \in \mathbb{R}$.

(b) If the covariates are the same across risks so that $X_1 = X_2 = X$, then for F_1 and F_2 the marginal CDFs, the set

$$\{\mu \mid F_1(\mu, \frac{\mu + a_1}{a_2} | \gamma) + b_1 F_2(\mu, \frac{\mu + a_1}{a_2} | \gamma)$$
$$= c_1 F_1(\frac{\mu + d_1}{d_2}, \frac{\mu + d_3}{d_4} | \gamma_o) + e_1 F_2(\frac{\mu + d_1}{d_2}, \frac{\mu + d_3}{d_4} | \gamma_o)\}.$$

has Lebesgue measure zero for $\gamma \neq \gamma_o$ and all constants $a, b, c, d, e \in \mathbb{R}$.

A2.2' - Covariates -

In addition to A2.2, assume the following: (iii) For some $t \in \{1, \ldots, T\}$ satisfying A2.2.(ii), $\exists\ h_j \in \{1, \ldots, k_j\}$, *s.t.* for j=1,2

(a) $\beta_{jh_j} \neq 0$,

(b) Conditional on $\tilde{X}_{jh_j}(t)$ defined as the vector of covariates with the h_j-th element of the t-th period deleted, $X_{jh_j}(t)$ has positive density in an open neighborhood in $\mathbb{R}$ *a.s.* P_X.

A2.3' - Parameters -

The true parameter vector $\theta_o = (\beta_{1o}, \beta_{2o}\tau_{1o}, \tau_{2o}, \gamma_o)$ is an interior point of the parameter space, the compact set $B_1 \times B_2 \times L_1 \times L_2 \times M$, where $B_1 \subset \mathbb{R}^{k_1}$, $B_2 \subset \mathbb{R}^{k_2}$, $L_1, L_2 \subset \mathbb{R}^T$, $M \subset \mathbb{R}^m$.

The additional assumptions on the distribution function and covariates are analagous to the additional assumptions made in the single risk case, and allow for separation of the γ parameter from the evaluation points. For the case where the covariates are the same, the distributional assumption is modified slightly to account for differences in the way that derivatives of the *CDFs* are taken. In practical terms however, the two assumptions are quite similar in spirit.

Lemma 2.4 - Given A2.1'-A2.3', the choice-specific interval probabilities are uniquely determined by choice of $\theta=(\beta_1,\beta_2,\tau_1,\tau_2,\gamma)$.

Proof: As before, I will restrict the discussion to the case where $\gamma \neq \gamma_o$. The notation that I use is a multiple risk generalization of the notation used in Lemma 1.3 above. First recall that the cause-specific evaluation points for period t are given by γ_t^j, j=1,2. For period s satisfying A2.2' above, define $\tilde{X}_1(s)'$, $\tilde{X}_2(s)'$, $\tilde{\beta}_1$, $\tilde{\beta}_2$ as the relevant vectors with the h_1 and h_2 elements deleted as required. The analogue to the single risk R-evaluation point functions is given by

$R_s^j(\theta, X, \Delta_j)$

$$= log \left[exp(\psi^j_{s-1}(\theta)) + exp(\tilde{X}_j(s)\tilde{\beta}_j + (X_{jh_j}(s) + \Delta_j)\tilde{\beta}_{jh_j}) \ \tau_j(s) \right]$$

for j=1,2. Note once again that $R_s^j(\theta, X, 0) = \psi^j_s(\theta)$.

The cause specific probabilities for the interval $(t-1, t]$, $P_t^j(\theta)$ are given by equations of the form (2.3) and (2.4). Manipulation of the integrals in (2.3) and (2.4) reveals that $\Sigma_j \ P_t^j(\theta) = F(\psi^1_t(\theta), \psi^2_t(\theta)) - F(\psi^1_{t-1}(\theta), \psi^2_{t-1}(\theta))$. Suppose that there exits a $\theta \neq \theta_o$ that generates the same set of choice-specific interval probabilities so that $P_t^j(\theta) = P_t^j(\theta_o)$ *a.s.* P_X, $\forall\ t$, j=1,2. In particular, for period *s* described above

$$F(\psi^1_s(\theta), \psi^2_s(\theta)) - F(\psi^1_{s-1}(\theta), \psi^2_{s-1}(\theta))$$

$$= \quad F(\psi^1_s(\theta_o), \psi^2_s(\theta_o)) - F(\psi^1_{s-1}(\theta_o), \psi^2_{s-1}(\theta_o))$$

a.s. P_X. Since ψ^1_{s-1} and ψ^2_{s-1} are invariant with respect to $X_1(s)$, this implies that

$$\lim_{\Delta_1, \Delta_2 \to 0} \frac{F(R_s^1(\theta, X, \Delta_1), R_s^2(\theta, X, \Delta_2) \mid \gamma) - F(\psi^1_s(\theta), \psi^2_s(\theta) \mid \gamma)}{\Delta_1 \ \ \Delta_2}$$

$$= \lim_{\Delta_1, \Delta_2 \to 0} \frac{F(R_s^1(\theta_o, X, \Delta_1), R_s^2(\theta_o, X, \Delta_2) \mid \gamma_o) - F(\psi^1_s(\theta_o), \psi^2_s(\theta_o) \mid \gamma_o)}{\Delta_1 \ \ \Delta_2}.$$

If the X's differ across risks, then taking the limits gives the condition that for all $X^1_{h_1}(s)$, $X^2_{h_2}(s)$ in an open neighborhood of $\mathbb{R}^2$,

$$f(\psi^1_s(\theta),\psi^2_s(\theta)\,|\,\gamma) = \mu_3\ f(\psi^1_s(\theta)+\mu_1,\psi^2_s(\theta)+\mu_2\,|\,\gamma_\circ)$$

where

$$\mu_j = \psi^j_s(\theta_\circ) - \psi^j_s(\theta), \quad j=1,2$$

$$\mu_3 = \frac{\overset{\circ}{\beta}_{1h_1} exp(X_1(s)\overset{\circ}{\beta}_1)\ \overset{\circ}{\tau}_1(s)}{\beta_{1h_1} exp(X_1(s)\beta_1)\ \tau_1(s)}\ \frac{exp(\psi^1_s(\theta))}{exp(\psi^1_s(\theta_\circ))}$$
$$\frac{\overset{\circ}{\beta}_{2h_2} exp(X_2(s)\overset{\circ}{\beta}_2)\ \overset{\circ}{\tau}_2(s)}{\beta_{2h_2} exp(X_2(s)\beta_2)\ \tau_2(s)}\ \frac{exp(\psi^2_s(\theta))}{exp(\psi^2_s(\theta_\circ))}.$$

Since by A2.2.(iii)' the X's have positive density in an open neighborhood, then the set of (ψ^1_s,ψ^2_s)-evaluation points satisfying this relationship has positive measure, violating A2.1.(ii)'.

If the X's are the same across risks so that $X_1 = X_2 = X$, then the proof must be modified. The existence of a θ which yields the same probabilities implies that for all $X_h(s)$ in an open neighborhood in $\mathbb{R}$,

$$\lim_{\Delta \to 0} \frac{F(R_s^1(\theta,X,\Delta),R_s^2(\theta,X,\Delta)\,|\gamma) - F(\psi_s^1(\theta),\psi_s^2(\theta)\,|\gamma)}{\Delta}$$

$$= \lim_{\Delta \to 0} \frac{F(R_s^1(\theta_o,X,\Delta),R_s^2(\theta_o,X,\Delta)\,|\gamma_o) - F(\psi_s^1(\theta_o),\psi_s^2(\theta_o)\,|\gamma_o)}{\Delta}.$$

Taking the limit,

$$F_1(\psi_1(\theta),\psi_2(\theta)\,|\gamma) + b_1\; F_2(\psi_1(\theta),\psi_2(\theta)\,|\gamma)$$
$$= d_1\; F_1(\psi_1(\theta_o),\psi_2(\theta_o)\,|\gamma_o) + e_1\; F_2(\psi_1(\theta_o),\psi_2(\theta_o)\,|\gamma_o)$$

where

$$b_1 = \frac{\beta_{2h_2} exp(X_2(s)\beta_2)\;\;\tau_2(s)}{\beta_{1h_1} \exp(X_1(s)\beta_1)\;\;\tau_1(s)}\;\frac{exp(\psi_s^2(\theta))}{exp(\psi_s^1(\theta))}$$

$$d_1 = \frac{\overset{o}{\beta}_{1h_1} exp(X_1(s)\overset{o}{\beta}_1)\;\;\overset{o}{\tau}_1(s)}{\beta_{1h_1} exp(X_1(s)\beta_1)\;\;\tau_1(s)}\;\frac{exp(\psi_s^1(\theta_o))}{exp(\psi_s^1(\theta))}$$

$$e_1 = \frac{\overset{o}{\beta}_{2h_2} exo(X_2(s)\overset{o}{\beta}_2)\;\;\overset{o}{\tau}_2(s)}{\beta_{1h_1} exp(X_1(s)\beta_1)\;\;\tau_1(s)}\;\frac{exp(\psi_s^2(\theta_o))}{exp(\psi_s^1(\theta))}.$$

By A2.2.(iii)', the set of points satisfying this relation has positive Lebesgue measure so that A2.1.(ii)' is violated.

Theorem 2.2 - Given A2.1'-A2.3', then almost surely a unique maximum likelihood estimator satisfying $\hat{\theta}_N \equiv (\hat{\beta}^1_N, \hat{\beta}^2_N, \hat{\tau}^1_N, \hat{\tau}^2_N, \hat{\gamma}_N)$, $\partial L_N(\hat{\theta})/\partial\theta = 0$ eventually exists and $\hat{\theta}_N \rightarrow \theta_o$. Furthermore, $\sqrt{N}(\hat{\theta} - \theta_o)$ $\overset{a}{\sim} N(0, J(\theta_o)^{-1})$.

Proof: A1.1'-A1.3' and the lemmas imply that all of the consistency and asymptotic normality conditions of McFadden Theorems 2 and 3 are satisfied.

3. *Conclusion*

In this chapter I present an extension of the Han and Hausman technique for the semi-parametric estimation of single and competing risk duration models to the important case where covariates vary over time. While semi-parametric estimators for models with time-varying covariates exist, they have previously been confined to the single risk framework. In this chapter, identification and asymptotic normality are demonstrated for the more general case.

In presenting a unified framework for single and competing risk models, HH demonstrate a correspondence between semi-parametric hazard models and existing discrete choice models. That result carries over to the time-varying case, which reduces to the HH case if variables are constant over time. The correspondence with discrete choice models provides considerable advantages since there is a large literature on the properties and extensions of discrete choice estimators which should provide directions for future research. For example, existing non-parametric techniques for the analysis of ordered discrete choice data may readily be applied.

The current discussion is restricted to demonstrating that the semi-parametric time-varying model can be estimated and has the standard maximum likelihood properties. The implementation of this model should be straightforward, and an empirical application of this technique is forthcoming. For computational reasons, some care should be taken in distinguishing between covariates which do and do not vary over time. For a model with a reasonable number of periods, the number of calculations increases significantly with the inclusion of time-varying covariates. The extra computational cost should be weighed against the possibility of misspecification resulting from treating covariates as constant over the entire duration.

Appendix 1 - Linear regression model for single-risk hazards with constant covariates.

The following transformation is well-known in the statistics literature: see for example, Kalbfleisch and Prentice [1980] who refer extensively to the regression format for the hazard model. For the true β vector and given realizations on X, the density function for the random variable for duration t is given by

$$f_t(t|X,\beta) = h(t)\ H(t), \tag{A1.1}$$

where

$$h(t) = \lambda_o(t)\ exp(X\beta),$$

$$H(t) = exp(-\textstyle\int_0^t h(s)\ ds).$$

Next, define a transformation on t to ϵ where ϵ is given by

$$\epsilon = g(t) = \textstyle\int_0^t h(s)\ ds\ , \tag{A1.2}$$

$$d\epsilon = h(t)\ dt.$$

Transforming the random variable t to ϵ yields:

(A1.3) $\quad f_\epsilon(\epsilon|X,\beta)$

$$= f_t(g^{-1}(\epsilon)\,|X,\beta)\ \ |J|$$

$$= f_t(t|X,\beta)\ \left|\frac{dt}{d\epsilon}\right|$$

$$= \quad h(t)\; exp(-g(t)) \left| h(t)^{-1} \right|$$

$$= \quad exp(-\epsilon)$$

which is recognizable as the density function for a unit mean exponential random variable.

A second change of variables produces the desired result. Define a natural log transformation of ϵ:

(A1.4)
$$\mu = g(\epsilon) = log(\epsilon)$$
$$g^{-1}(\mu) = exp(\mu)$$
$$\left| \frac{dg^{-1}(\mu)}{d\mu} \right| = exp(\mu).$$

Thus, transforming ϵ to μ yields the density function

(A1.5) $f_\mu(\mu|X,\beta)$

$$= \quad f_\epsilon(exp(\mu)\,|X,\beta)\;\; |J|$$

$$= \quad exp(-exp(\mu)) \left| \frac{dg^{-1}(\mu)}{d\mu} \right|$$

$$= \quad exp(-exp(\mu))\; exp(\mu)$$

$$= \quad exp(\mu - exp(\mu))$$

where the random variable μ can be recognized as having an extreme value distribution $F_\mu(\mu) = 1 - exp(-exp(\mu))$.

Using the definitions of ϵ and μ yields

(A1.6) $$\begin{aligned}\mu &= log(\epsilon) \\ &= log(\textstyle\int_0^t h(s)\ ds) \\ &= log(\textstyle\int_0^t \lambda_o(s)\ ds) + X\beta,\end{aligned}$$

and it follows immediately that

(A1.7) $$log(\textstyle\int_0^t \lambda_o(s)\ ds) = -X\beta + \mu,$$

where μ has an extreme value distribution as defined above.

Appendix 2 - Linear regression model for proportional hazards, non-time varying variables, with gamma heterogeneity.[1]

Consider first the case where the covariates are assumed to be constant over time. The linear specification for the proportional hazards model defined previously is

$$\text{(A2.1)} \qquad \log \int_0^t \lambda_o(s)\ ds = -X\beta + \epsilon$$

where ϵ has an extreme value distribution. Then taking exponentials,

$$\text{(A2.2)} \qquad H(t)\ \exp(X\beta) = \exp(\epsilon)$$

where $H(t) = \int_0^t \lambda_o(s)\ ds$ is the integrated hazard function.

Now modify (A2.2) to allow for gamma heterogeneity by adding the multiplicative error term $\exp(\omega)$

$$\text{(A2.3)} \qquad H(t)\ \exp(X\beta) = \exp(\epsilon + \omega)\ .$$

This implies that the relationship given in (A2.1) holds with ϵ replaced by the error term $(\epsilon + \omega)$. In this context, ω corresponds to the logarithm of the gamma variable defined by Lancaster [1979].

---

[1] I am indebted to Aaron Han for showing me the form of this derivation in the time-constant case.

The estimation of this model requires knowledge of the distribution of the new error term. Define $I(t) = H(t)\ exp(X\beta)$ and $\xi = exp(\epsilon + \omega)$, then rewrite (A2.3) as

$$I(t) = \xi \ . \tag{A2.4}$$

Since $I(t)$ is monotonically increasing in t, the probability that the failure time is greater than some value t is simply the integral of the density of ξ over the portion of the support of ξ that is greater than $I(t)$. Using the form for the survivor function derived by Lancaster [1979] and taking advantage of the correspondence between models:

$$\int_{I(t)}^{\infty} f_{\xi}(s)\ ds = (1 + \sigma^2\ I(t))^{-1/\sigma^2} \tag{A2.5}$$

where σ^2 is the gamma variance. Differentiating both sides with respect to $I(t)$ gives

$$f_{\xi}(I(t)) = (1 + \sigma^2\ I(t))^{-(1/\sigma^2+1)} . \tag{A2.6}$$

Finally, applying a change of variables from ξ back to $\exp(\epsilon + \omega)$ gives the result, for $\gamma = \epsilon + \omega = log(I(t))$

(A2.7) $$f_\gamma(\gamma) = (1 + \sigma^2 \; exp \; \gamma)^{-(1/\sigma^2+1)} \; exp(\gamma)$$

so that the equivalent, regression form of the model (A2.1) with gamma heterogeneity is

(A2.8) $$log \int_0^t \lambda_o(s) \; ds = -X\beta + \gamma,$$

where the density function of γ is given in (A2.7).

For the case where the covariates vary over time, slightly different notation is needed. The regression form of the model with heterogeneity is given by

(A2.9) $$log \left(\sum_{i=1}^{t} exp(X(i)\beta) \int_{i-1}^{i} \lambda_o(s) \; ds \right) = \epsilon + \omega,$$

(A2.10) $$R(t) = exp(\epsilon + \omega) = \xi.$$

where the $R(t)$ is analagous to the $I(t)$ in (A2.4). Thus, by arguments similar to those given above, the density funciton of $\gamma = \epsilon + \omega = log(R(t))$ is given by (A2.7). The regression form of the time-varying covariate model is then

(A2.11) $$log \left(\sum_{i=1}^{t} exp(X(i)\beta) \int_{i-1}^{i} \lambda_o(s) \; ds \right) = \gamma,$$

where γ has the density function given in (A2.7).

Appendix 3 - Limits of integration for the bivariate competing risks model with constant covariates.

This derivation also appears in Han and Hausman [1986] and is the special, non-time varying case of the derivation in chapter 3 of this thesis. The probability that a failure observed during the interval $(t-1,t]$ is of type 1 is given by the likelihood defined in the text

$$\text{(A3.1)} \qquad H_1(t) = \int_{I^1_{t-1}+X\beta}^{I^1_t + X\beta} \int_{g(\epsilon_1)}^{\infty} f(\epsilon_1 . \epsilon_2) \; d\epsilon_2 \, d\epsilon_1$$

where $g(\epsilon_1)$ is a function relating the error term ϵ_2 to ϵ_1 so that the latent failure time of risk 2 is equal to the failure time for risk 1 implied by the realization of ϵ_1.

To solve for the function $g(\cdot)$, first recall the linear specification for the two hazards

$$\text{(A3.2)} \qquad \begin{aligned} I^1_t &= -X_1\beta_1 + \epsilon_1 \ , \\ I^2_t &= -X_2\beta_2 + \epsilon_2 \ . \end{aligned}$$

Since a failure of type 1 is observed at time t, the function $g(\cdot)$ will be evaluated at all ϵ^* in the interval $(t-1,t]$. Following Han and Hausman [1986], I assume linearity of the hazards for both risks over discrete intervals. The regression form for the hazards given in (A3.2) can then be written, for arbitrary ϵ^* lying in the interval $(t-1,t]$, as

$$I^1_{t-1} = \kappa(I^1_t - I^1_{t-1}) = -X_1\beta_1 + \epsilon_1 ,$$
$$I^2_{t-1} = \kappa(I^2_t - I^2_{t-1}) = -X_2\beta_2 + \epsilon_2 . \tag{A3.3}$$

where $\kappa = (\epsilon - \epsilon^*_{t-1})/(\epsilon_t - \epsilon_{t-1})$ gives the fraction of the interval from $(\epsilon_{t-1},\epsilon_t]$ that ϵ^* is greater that ϵ_{t-1}. Solving for ϵ

$$\kappa = \frac{\epsilon^* - (I^1_{t-1} + X\beta)}{I^1_t - I^1_{t-1}} . \tag{A3.3}$$

Applying the proportionality factor κ to the second equality in (A3.3) and then solving for ϵ yields the function given in the text,

$$g(\epsilon_1) = I^2_t + X_2\beta_2 + \frac{I^2_t - I^2_{t-1}}{I^1_t - I^1_{t-1}}(\epsilon_1 - (I_{t-1} + X_1\beta_1)) . \tag{A3.4}$$

References

Aaron, Henry J. [1984]. Economic Effects of Social Security, Washington, D.C.: The Brookings Institution.

Amemiya, Takeshi [1981]. "Qualitative Response Models: A Survey," Journal of Economic Literature, 19, pp. 1483-1536.

Baily, Martin N. [1977]. "On the Theory of Lay-offs and Unemployment," Econometrica, 45, pp. 1043-64.

Berndt, E. K. and B. H. Hall, R. E. Hall, J. A. Hausman [1974]. "Estimation and Inference in Nonlinear Structural Models," Annals of Economic and Social Measurement, 3/4, pp. 653-65.

Blinder, Alan S., Gordon, Roger H., and Donald Wise [1980]. Reconsidering the Work Disincentive Effects of Social Security," National Tax Journal, 34, pp. 467-72.

Boskin, Michael J. [1977]. "Social Security and Retirement Decisions," Economic Inquiry, 15, pp. 1-25.

_______ and Michael D. Hurd [1978]. "The Effect of Social Security on Early Retirement," Journal of Public Economics, 10, pp.361-77.

Burdett, Kenneth, and Dale T. Mortensen [1980]. "Search Layoffs and Market Equilibrium," Journal of Political Economy, 88, pp. 652-72.

Burkhauser, Richard V. [1979]. "The Pension Acceptance Decision of Older Workers," Journal of Human Resources, 33, pp. 470-83.

Burtless, Gary [1986]. "Social Security, Unanticipated Benefit Increases and the Timing of Retirement," Review of Economic Studies, 53, pp. 781-805.

_______ and Robert Moffitt [1984]. "The Effect of Social Security Benefits on the Labor Supply of the Aged," in Henry J. Aaron and Gary Burtless, eds. Retirement and Economic Behavior, Washington, D.C.: The Brookings Institution.

Clark, Kim B., and Lawrence H. Summers [1979]. "Labor Market Dynamics and Unemployment: A Reconsideration," Brookings Papers on Economic Activity, 1979:1, pp. 13-60.

Cox, D.R. [1972]. "Regression Models and life tables," Journal of the Royal Statistical Society, B, 34, pp. 187-220.

_______ and D. Oakes [1984]. Analysis of Survival Data, London: Chapman and Hall.

Diamond, Peter A. [1977]. "A Framework for Social Security Analysis," Journal of Public Economics, 8, pp. 275-98.

_______ and Jerry A. Hausman [1984]. "The Retirement and Unemployment Behavior of Older Men," in Henry J. Aaron and Gary Burtless, eds. Retirement and Economic Behavior, Washington, D.C.: The Brookings Institution.

Feldman, Robert A. [1985]. "An Examination of Health Status Measurement in Empirical Models of Retirement Behavior," unpublished.

Feldstein, Martin S. [1978]. "The Effect of Unemployment Insurance on Temporary Layoff Unemployment," American Economic Review, 68, pp. 834-46.

_______ [1974]. "Social Security, Induced Retirement, and Aggregate Capital Accumulation," Journal of Political Economy, 92, pp. 905-26.

Fields, Gary S., and Olivia S. Mitchell [1984]. Retirement, Pensions and Social Security, Cambridge: MIT Press.

Fox, Alan [1976]. "Alternative Means of Earnings Replacement Rates for Social Security Benefits," from Reaching Retirement Age, Social Security Administration Report 47, Office of Research and Statistics: GPO.

Gordon, Roger H., and Alan S. Blinder [1980]. "Market Wages, Reservation Wages and Retirement," Journal of Public Economics, 14, pp. 277-308.

Gustman, Alan L., and Thomas L. Steinmeier [1986]. "A Structural Retirement Model," Econometrica, 54, pp. 555-584.

_______ [1984]. "Partial Retirement and the Analysis of Retirement Behavior," Industrial and Labor Relations Review, 37, pp. 403-415.

_______ [1983]. "Retirement Flows," NBER working paper, no. 1069.

Hall, Robert E. [1972]. "Turnover in the Labor Force," Brookings Papers on Economic Activity, 3:1972, pp. 709-56.

Han, Aaron K., and Jerry A. Hausman [1986]. "Semiparametric Estimation of Duration and Competing Risk Models," unpublished MIT working paper.

Hausman, Jerry A. [1978]. "Specification Tests in Econometrics," Econometrica, 46, pp. 1251-1272.

Hausman, Jerry A., and David A. Wise [1985]. "Social Security, Health Status, and Retirement," in David A. Wise, ed. Pensions, Labor, and Individual Choice, Chicago: University of Chicago Press.

_______ [1978]. "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," Econometrica, 46, pp. 406-426.

Heckman, James J., and Burton Singer [1986]. "Econometric Analysis of Longitudinal Data," in Griliches, Z. and M. Intriligator, eds., The Handbook of Econometrics, vol 3., Amsterdam: North Holland.

_______[1984a]. "Econometric Duration Analysis," Journal of Econometrics, 24, pp. 63-132.

_______ [1984b]. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," Econometrica, 52, pp. 271-320.

Hurd, Michael D., and Michael J. Boskin [1984]. "The Effect of Social Security on Retirement in the Early 1970s," Quarterly Journal of Economics, pp. 767-90.

Kaitz, Hyman B. [1970]. "Analyzing the Length of Spells of Unemployment," Monthly Labor Review, 93, pp. 11-20.

Kalbfleisch, J.D., and Prentice, R.L. [1980]. The Statistical Analysis of Failure Time Data, New York: John Wiley and Sons.

Kaplan, E.L., and P. Meier [1958]. "Nonparametric Estimation from Incomplete Observations," Journal of the American Statistical Association, 53, pp. 457-481.

Katz, Lawrence [1986]. "Layoffs, Recall and the Duration of Unemployment," NBER working paper, no. 1825.

_______ [1985]. "Worker Mobility and Unemployment," Ph.D. Dissertation, MIT.

Kiefer, Nicholas M., Lundberg, Shelly J., and George R. Neumann [1984]. "How Long is a Spell of Unemployment?: Illusions and Biases in the Use of CPS Data." NBER working paper,, no. 1467.

Lancaster, Tony [1979]. "Econometric Methods for the Duration of Unemployment," Econometrica, 47, pp. 939-56.

McCall, John J. [1970]. "Economics of Information and Job Search," Quarterly Journal of Economics, 84, pp. 113-26.

McFadden, Daniel [1986]. "A Method of Simulated Moments for Estimation of Multinomial Probits without Numerical Integration," unpublished.

________ [1984]. "Qualitative Response Models", in Grilichis, Z. and M. Intriligator, eds., The Handbook of Econometrics, vol. 2, Amsterdam: North Holland.

Manton, Kenneth G., and Eric Stallard, and James W. Vaupel [1986]. "Alternative Models for the Heterogeneity of Mortality Risks Among the Aged," Journal of the American Statistical Association, 81, pp. 635-644.

Marston, Stephen T. [1975]. "The Impact of Unemployment Insurance on Job Search," Brookings Papers on Economic Activity, 1:1975, pp. 13-48.

Meyer, Bruce D. [1987]. "Unemployment Insurance and Unemployment Spells," unpublished.

Moffitt, Robert [1985]. "Unemployment Insurance and the Distribution of Unemployment Spells," Journal of Econometrics, 38, pp. 85-101.

Mortensen, Dale T. [1977]. "Unemployment Insurance and Job Search Decisions," Industrial and Labor Relations Review, 30, pp. 505-17.

Paquette, Lynn [1984]. "Liquidity Constraints and the Consumption Patterns of the Elderly," unpublished.

Parnes, Herbert S. [1981]. "Chapter 8: Summary and Conclusions," Work and Retirement, Cambridge, MIT Press,

_______ and Gilbert Nestel [1981]. "The Retirement Experience," in Herbert S. Parnes, ed., Work and Retirement, Cambridge, MIT Press.

Pissarides, Christopher A. [1982]. "Job Search and the Duration of Layoff Unemployment," Quarterly Journal of Economics, 97, pp. 543-67.

Poterba, James M., and Lawrence H. Summers [1986]. "Reporting Errors and Labor Market Dynamics," Econometrica, 54, pp. 1319-38.

Pratt, John W. [1981]. "Concavity of the Log Likelihood," Journal of the Americal Statistical Association, 76, pp. 103-06.

Prentice, R.L., and L.A. Gloeckler [1978]. "Regression Analysis of Grouped Survival Data with Applications to Breast Cancer Data," Biometrics, 34, pp. 57-67.

Prentice, R.L., Kalbfleisch, J.D., et. al. [1978]. "The Analysis of Failure Times in the Presence of Competing Risks," Biometrics, 34, pp. 541-54.

Rao, C. R. [1965]. Linear Statistical Inference and Its Applications, N.Y.: John Wiley and Sons.

Salant, Stephen W. [1977]. "Search Theory and Duration Data: A Theory of Sorts," Quarterly Journal of Economics, 91, pp. 39-57.

Topel, Robert [1983]. "On Layoffs and Unemployment Insurance," The American Economic Review, 73, pp. 541-559.

U.S. Bureau of the Census [various years]. Survey of Income and Program Participation (SIPP), Waves 1-4 Relational Core Microdata File Technical Documentation, Washington D.C.: The Bureau.

Zabalza, A., Pissarides, C., and M. Barton [1980]. "Social Security and the Choice Between Full-time Work, Part-time Work and Retirement," Journal of Public Economics, 14, pp. 245-76.